On the non-local geometry of turbulence

Thesis by

Iván Bermejo-Moreno

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2008

(Defended April 18, 2008)

To my family

and to the memory of Pawel Buraczewski

Acknowledgments

This Ph.D. and the marvelous experience that I lived through it were possible thanks to the help of many. First, I want to thank my advisor, Professor Dale Pullin, for his guidance, encouragement, patience, and support. His ability to keep always a larger view both in the course of research and also in life has given another dimension to my learning experience. Being his student has been an honor and a pleasure. I would like to thank the members of my thesis committee, Prof. Beverley McKeon, Prof. Daniel Meiron, and Prof. Joseph Shepherd for their valuable time and for showing interest in this research.

Interacting with the GALCIT community has been very rewarding. Particularly, I want to thank my friends from the former Iris Lab. Also, in the Aeronautics department, I would like to thank Alexandra Katsas and Dimity Nelson for their invaluable help. I am grateful to the members of the Caltech faculty, both in GALCIT and in other departments, who were involved in my education. Also at Caltech, I am thankful to Marjory Gooding, Jim Endrizzi, and Athena Trentin, for their outstanding dedication at the International Student Programs. I want to thank Dr. Paul O'Gorman, at Caltech, and Dr. Kiyosi Horiuti, at the Tokyo Institute of Technology, for providing the numerical turbulence databases used in this work, and Prof. Emmanuel Càndes and Dr. Lexing Ying for valuable discussions regarding curvelets.

The Fulbright Program gave me the opportunity to come to the U.S.A., and in particular to Caltech, to pursue a Ph.D. degree. With it, they gave me also one of the most wonderful experiences of my life, for which I will always be grateful. I wish to thank especially George L. Argyros, for his trust in me and his support, and Ann Z. Kerr, for her friendship and for being an example and inspiration. I feel honored and very fortunate to have met her.

I would like to thank my friends, those in the distance and those I met here, for making this journey so special. Sharing this time with them has been the best way I could have ever imagined to discover the amazing place California is.

My deepest gratitude goes to my parents, Atanasio and Ana Elisa Aura, and to my sister Sonia. Thanks to them, this dream came true.

This work has been supported in part by the National Science Foundation under grants DMS-0353838 and DMS-0714050, and by the Fulbright Program, under sponsorship of the U.S. Department of State.

Abstract

A multi-scale methodology for the study of the non-local geometry of eddy structures in turbulence is developed. Starting from a given three-dimensional field, this consists of three main steps: extraction, characterization, and classification of structures. The extraction step is done in two stages: first, a multi-scale decomposition based on the curvelet transform is applied to the full three-dimensional field, resulting in a finite set of component fields, one per scale; second, by iso-contouring each component field at one or more iso-contour levels, a set of closed iso-surfaces is obtained that represents the structures at that scale. For periodic domains, those structures intersecting boundaries are reconnected with their continuation in the opposite boundaries. The characterization stage is based on the joint probability density function (jpdf), in terms of area coverage on each individual iso-surface, of two differential-geometry properties—the shape index and curvedness—plus the stretching parameter, a dimensionless global invariant of the surface. Taken together, this defines the geometrical signature of the iso-surface. The classification step is based on the construction of a finite set of parameters, obtained from algebraic functions of moments of the jpdf of each structure, that specify its location as a point in a multi-dimensional 'feature space'. At each scale the set of points in feature space represents all structures at that scale, for the specified iso-contour value. This allows the application, to the set, of clustering techniques that search for groups of structures with a common geometry.

Results are presented of a first application of this technique to a passive scalar field obtained from 512^3 direct numerical simulation of scalar mixing by forced, isotropic turbulence ($Re_\lambda = 265$). These show transition, with decreasing scale, from blob-like structures in the larger scales to blob- and tube-like structures with small or moderate stretching in the inertial range of scales, and then

toward tube and predominantly sheet-like structures with high level of stretching in the dissipation range of scales. Implications of these results for the dynamical behavior of passive scalar stirring and mixing by turbulence are discussed.

We apply the same methodology to the enstrophy and kinetic energy dissipation rate instantaneous fields of a second numerical database of incompressible homogeneous isotropic turbulence decaying in time obtained by DNS in a periodic box. Three different resolutions are considered: 256^3, 512^3, and 1024^3 grid points—with $k_{\max}\bar{\eta}$ approximately 1, 2, and 4, respectively, the same initial conditions and $Re_\lambda \approx 77$. This allows a comparison of the geometry of the structures obtained for different resolutions. For the highest resolution, structures of enstrophy and dissipation evolve in a continuous distribution from blob-like and moderately stretched tube-like shapes at the large scales to highly stretched sheet-like structures at the small scales. The intermediate scales show a predominance of tube-like structures for both fields, much more pronounced for the enstrophy field. The dissipation field shows a tendency toward structures with lower curvedness than those of the enstrophy for intermediate and small scales. The 256^3 grid resolution case ($k_{\max}\bar{\eta} \approx 1$) was unable to detect the predominance of highly stretched sheet-like structures at the smaller scales.

The same methodology, but without the multi-scale decomposition, is then applied to two scalar fields used by existing local criteria for the eduction of tube- and sheet-like structures in turbulence, Q and $[A_{ij}]_+$, respectively, obtained from invariants of the velocity gradient tensor and alike in the 1024^3 case. This adds the non-local geometrical characterization and classification to those local criteria, assessing their validity in educing particular geometries.

Finally we introduce a new methodology for the study of proximity issues among different sets of structures, based also on geometrical and non-local analyses. We apply it to four of the fields previously studied. Tube-like structures of Q are mainly surrounded by sheets of $[A_{ij}]_+$, which appear at close distances. For the enstrophy, tube-like structures at an intermediate scale are primarily surrounded by sheets of smaller scales of the enstrophy and structures of dissipation at the same and smaller scales. A secondary contribution results from tubes of enstrophy at smaller scales appearing at farther distances. Different configurations of composite structures are presented.

Contents

List of Figures

List of Tables

Nomenclature

Greek letters

α_ℓ, β_ℓ	Angles defining the center slope of the frequency wedge (curvelets)	12
γ_k	Array of parameters stored in the conditional array map (CAM)	71
ϵ	Local dissipation	45
ϵ_{ijk}	Levi-Civita symbol	45
$\bar{\eta}$	Average Kolmogorov length scale	30
ϑ	Eigenvalues of A_{ij}	63
κ_1	Maximum curvature	15
κ_2	Minimum curvature	15
κ_n	Normal curvature	16
Λ	Curvedness	15
λ	Stretching parameter	19
μ	Characteristic length scale of a closed surface	18
μ_c	Mean passive scalar gradient magnitude in x_1 direction	6
ν	Kinematic viscosity	44
ξ, ζ	Local properties mapped onto a surface	72
ρ	Density	63
ϱ	Radial polar coordinate in the plane of principal curvatures	16
σ_l	Local scaling parameter	22
Υ	Shape index	15

ϕ	Azimuthal polar coordinate in the plane of principal curvatures	16
$\varphi^D_{j,\ell,k}$	Curvelets in physical space	11
$\hat{\varphi}^D_{j,\ell,k}$	Curvelets in Fourier space	11
Ω_{ij}	Rotation-rate tensor	45
ω	Wavenumber (Fourier domain)	11
ω_i	Vorticity field ($i = 1, 2, 3$)	45
$\omega_i\omega_i$	Local enstrophy	45

Roman letters

$\hat{\boldsymbol{A}}$	Locally scaled affinity matrix	22
A	Area of a surface	18
$\boldsymbol{a}$	Vector contained on the tangent plane at a point P of a surface	16
a_i	Average distance from an element to other elements of its same cluster (in the clustering algorithm)	26
a_i	Element i of the set $\mathcal{A}$ (in the structure interaction analysis)	71
A_{ij}	Symmetric second-order tensor $S_{ik}\Omega_{kj} + S_{jk}\Omega_{ki}$	3
$[A_{ij}]_+$	Largest remaining eigenvalue of A_{ij} after removing $[A_{ij}]_\omega$	63
$[A_{ij}]_-$	Smallest remaining eigenvalue of A_{ij} after removing $[A_{ij}]_\omega$	63
$[A_{ij}]_\omega$	Eigenvalue of A_{ij} associated with the eigenvector most aligned with the vorticity field, ω_i	63
b_i	Average distance from an element to the elements in the closest cluster	26
C	Dimensionless curvedness	18
$\hat{C}$	Feature dimensionless curvedness center	24
c	Passive scalar	6
c'	Passive scalar fluctuation	6
c_j	Element j of the set $\mathcal{C}$ (in the structure interaction analysis)	71

$c^D(j,l,k)$	Curvelet coefficients	11
$\boldsymbol{D}$	Diagonalizing matrix	22
D	Diffusivity	6
d	Distance	72
D_{a_i}	Non-dimensionalizing length scale of the structure a_i	72
d_{ij}	Distance matrix	21
d_l	Lower distance of a probability density function	24
d_u	Upper distance of a probability density function	24
E	Set of elements, e_i, to cluster, $i = 1, ..., N$	21
$E(k)$	Energy spectrum of original field (containing all scales)	33
$E_i(k)$	Energy spectrum associated with component field at scale number i	33
F	Distance function in the space of parameters	22
f_k	Cumulative marginal probability density function of proximity for group indices g from 1 to k	78
g	Group index	72
G_g	Groups contained in set $\mathcal{B}$	72
j	Scale number (curvelets)	11
j_0	Minimum scale number (curvelets)	11
j_e	Maximum scale number (curvelets)	11
k	Spatial location index, $\{k_i, i = 1, 2, 3\}$ (curvelets)	11
$k_{\max}$	Largest dynamically significant wavenumber	30
$\boldsymbol{L}$	Normalized locally scaled affinity matrix	22
ℓ	Orientation index (curvelets)	11
L_i, L_i'	Integral length scales of component field at scale number i	33
$\boldsymbol{N}$	Normal vector to the tangent plane at a point P of a surface	16
N	Number of elements to cluster	21
n	Grid size of side cubic domain	11

N_g	Number of G_g groups contained in set $\mathcal{B}$	72
N_P	Number of parameters defining the feature space	21
p	Pressure field (in the equations of fluid mechanics)	63
p	Proximity (in the structure interaction analysis)	72
$\mathcal{P}$	Area-based joint probability density function	19
$\mathcal{P}_\mathcal{C}$	Area-based (marginal) probability density function of C	19
$\mathcal{P}_\mathcal{S}$	Area-based (marginal) probability density function of S	19
$\vec{\mathcal{PI}}(\xi,\zeta;p)$	Area-based joint probability density function in terms of the local properties (ξ,ζ) with averaged intensity component in terms of the local property p	73
$p[k]$	Parameters of feature space	21
R	Radius of curvature	16
r	Number of closest neighbors for local scaling (in the clustering algorithm)	22
Re	Reynolds number	42
Re_λ	Taylor Reynolds number	30
S	Absolute value of the shape index	16
$\hat{S}$	Feature absolute value of the shape index center	24
SC	Silhouette coefficient	26
Sc	Schmidt number	30
S_{ij}	Strain-rate tensor	45
$S_{ij}S_{ij}$	Local dissipation renormalized by $(2\,\nu)^{-1}$	45
t	Characteristic thickness of a sheet-like structure	76
t	Time variable (in the equations of fluid mechanics)	6
$\boldsymbol{u}$	Velocity vector field (with components $u_j, j = 1,2,3$)	6
$\overline{u^2}$	Characteristic squared integral velocity of original field	33

$\overline{u_i^2}$	Characteristic squared integral velocity of component field at scale number i	33
$\tilde{U}_{j,\ell}$	Frequency window (curvelets)	11
V	Volume inside a closed surface	18
$\tilde{V}_{j,\ell}$	Angular frequency window (curvelets)	11
$\tilde{W}_j$	Radial frequency window (curvelets)	11
$\boldsymbol{X}$	Matrix of eigenvectors	22
$\boldsymbol{x}$	Position vector (spatial coordinates $x_j, j = 1, 2, 3$)	6
$\mathcal{X}(\alpha)$	Set of extracted structures from a three-dimensional scalar field α	75
$\boldsymbol{Y}$	Renormalized matrix of eigenvectors	23

Acronyms

BIC	Bayesian information criterion	26
CAM	Conditional array map	71
DNS	Direct numerical simulation(s)	1
HSB	Hue-saturation-brilliance color space	76
jpdf	Joint probability density function	19
jpdf+i	Joint probability density function with intensity component	73
LES	Large eddy simulation(s)	2
MDM	Minimum distance map	71
pdf	Probability density function	13

Chapter 1

Introduction

1.1 The role of geometry

Observation of natural fluid flows indicates the presence of structures with apparent repeating geometries. Vortical structures in multiphase flows are commonly observed. The roll-up of an ocean wave before it breaks (pressure- and gravity-driven flow), the swirling motion of a hurricane around its center (pressure-driven flow affected by Coriolis and friction forces) or that of the stellar gas accretion disk occurring during the formation of galaxies (gravitational-driven flow), and the Kelvin–Helmholtz wave clouds formed between two layers of air of different density and speeds (shear-driven flow) are just a few examples. Turbulent fluid flows are no exception, often adding levels of complexity to the structure geometry owing to the multiple scales that comprise such flows.

Visualization experiments have provided means for the systematic study of geometrical structures in fluid flows and have substantially increased the number of known flows where repeating geometrical patterns are present. Experimental study of the flow past cylinders and spheres led to the discovery of the Kármán vortex street while experiments in turbulent mixing layers resulted in an exhaustive study of 'coherent' vortical structures (see Brown & Roshko, 1974). This work has stimulated theoretical analysis of pattern formation, for example, the description of eddying motions and flow patterns based on critical-point theory (see Perry & Chong, 1987).

Direct numerical simulations (DNS) have also proven to be a valuable tool in the search for geometrical structures in fluid flows. The organized cylindrical elongated vortices (so-called 'worms')

found in the intense vorticity of isotropic turbulence (Siggia, 1981; Kerr, 1985; Jiménez et al., 1993) are one remarkable example. The 'worms', however, remain a puzzle; their contribution to the kinetic energy dissipation is almost negligible and their role in turbulence dynamics remains an open question.

Structures in turbulent flows can be considered a consequence of the forces and boundary conditions driving the flow, but also can be seen as themselves producing some intrinsic properties of the turbulence. In the multi-scale ansatz based on self-similarity and the idea of energy cascade (Richardson, 1922; Kolmogorov, 1941*a*,*b*; Onsager, 1945), the external forces and the boundary conditions affect mainly large energy-containing scales, with diminished influence on progressively smaller eddies. The energy-containing scales then depend strongly on the external forces and boundary conditions and are not expected to be universal, while small-scale structures may be related to universal properties of turbulence, thereby exhibiting a generic geometric signature that may be characteristic of efficient cascade dynamics.

A geometrical characterization of those structures could provide improved understanding of cascade mechanics and dissipation-range dynamics, contributing potentially to the development of structure-based models of turbulence fine scales (see Townsend, 1951; Tennekes, 1968; Lundgren, 1982; Pullin & Saffman, 1993), subgrid-scale models for large-eddy simulation LES (see Misra & Pullin, 1997), and simulation methods based on multi-resolution decomposition by means of the wavelet transform (see Farge, 1992; Meneveau, 1991; Farge et al., 1996, 1999). Further, a better understanding of eddy structure at large Reynolds number may provide important insight into possibly singular or near-singular structures in the dynamics of the Euler equations (see Hou & Li, 2006) by elucidating the geometrical characterization of sites within the turbulent field where extreme dissipative or vortical events occur, and which are candidates for singularity formation in the limit of vanishing viscosity.

1.2 Previous identification criteria

Prior work on the identification of structures in turbulence addresses mainly the identification of vortex tube- and sheet-like structures with emphasis on vortex tubes. But the importance of sheet-like structures, where significant turbulent kinetic dissipation may be concentrated owing to their high amplitude of strain rate, and which may produce tubes by roll-up instabilities, has led to renewed interest in sheets. Most identification methods either for tubes, sheets, or both, are based on local measures of scalar fields obtained from the velocity-gradient tensor and/or the pressure field. They rely on physical aspects associated with a particular kind of structure either of turbulent flows or of simpler solutions of the Navier-Stokes equations (e.g., Burgers vortex tubes and sheets), whose phenomenology is extrapolated to turbulence.

Chong et al. (1990) classified regions with complex eigenvalues of the velocity-gradient tensor as vortex tubes (since the local streamlines are then closed or spiral in a reference frame moving with the fluid). The second-order invariant, Q, of the velocity-gradient tensor was used by Hunt et al. (1988), to define a vortex tube as the region with a positive value of Q, and the condition of a pressure lower than the ambient, while Ashurst et al. (1987) based their identification criterion on the sign of the intermediate eigenvalue of the strain-rate tensor, S_{ij}. Tanaka & Kida (1993) extended the identification criterion based on Q for the extraction of both tubes and sheets. Jeong & Hussain (1995) proposed a method based on the second largest eigenvalue, λ_2, of the tensor L_{ij} formed by summing the products of the symmetric, S_{ij}, and antisymmetric, Ω_{ij}, parts of the velocity-gradient tensor with themselves, $L_{ij} = S_{ik}S_{kj} + \Omega_{ik}\Omega_{kj}$. They define a vortex core as the region where λ_2 is negative. Horiuti (2001) combined this methodology with the physical explanations of the alignment of vorticity and the eigenvector associated with the intermediate eigenvalue of S_{ij} (Andreotti, 1997) to develop a new method in which the eigenvalues and eigenvectors of L_{ij} are reordered based on their alignment with the vorticity; then, regions are classified into vortex tubes, and so-called flat vortex sheets and curved vortex sheets depending on the relations of those reordered eigenvalues. Horiuti & Takagi (2005) proposed an improved method for the eduction of vortex sheet structures, based on local values of the largest eigenvalue of the tensor $A_{ij} = S_{ik}\Omega_{kj} + S_{jk}\Omega_{ki}$, once the eigenvalue

corresponding to the eigenvector maximally aligned with the vorticity is removed.

Based solely on the pressure field, Miura & Kida (1997) developed a methodology for extracting axes of tubular vortices as the loci of sectionally local minima of the pressure field (obtained by means of the sign of the second largest eigenvalue of the pressure Hessian evaluated at each point; positive values indicate pressure minima).

The majority of existing methods of identification are local, based on pointwise quantities used to discriminate whether each point belongs to one type of structure or another (or none). Regions of points sharing a common identity based on the local criterion applied can then be formed, but often that local analysis is the end of the identification process. Visualization of such regions has proved a helpful tool in its analysis, but here we seek a more automated, systematic approach to structure characterization.

Some non-local methods exist in the fluid mechanics literature. These classify structures considering their spatial extent and can handle a broader range of geometries. While local methods are often based on a priori physical knowledge of the particular geometry to be educed, non-local methods generally draw physical conclusions a posteriori, based on geometrical characteristics obtained from the educed structures. For example, an extended structural and fractal description of turbulence was proposed by Moisy & Jiménez (2004), who applied a box-counting method to sets of points of intense vorticity and strain-rate magnitude (educed by thresholding). They also analyzed geometrically individual structures, defined as a connected set of points satisfying the threshold criterion (thus, considering the spatial extent of such structures), based on their volume and spatial distribution, finding that intense vorticity and dissipation structures are concentrated in clusters of inertial size. Wang & Peters (2006) defined extended dissipation elements as the ensemble of grid cells from which the same pair of extremal points of the scalar field can be reached, and studied their characteristic linear distances.

1.3 Non-local, multi-scale, and clustering features

Our approach is based on a non-local, multi-scale methodology for the extraction, characterization, and classification of structures in turbulence from a geometrical perspective. It is non-local, focusing on the spatial extent of structures. The multi-scale analysis is performed through the curvelet transform, a higher-dimensional generalization of the wavelet transform. Presently, the structures are defined as iso-surfaces extracted, at different scales, from a three-dimensional scalar field obtained from a turbulent flow. The characterization and classification steps are based on measures of the geometry of iso-surfaces. The problem of shape analysis of free-form surfaces has been widely studied in the fields of computer graphics, computer vision and image understanding, (see Campbell & Flynn, 2001; Iyer et al., 2005; Dorai & Jain, 1997; Osada et al., 2001; Zaharia & Prêteux, 2001). Our method characterizes each individual structure in terms of local differential-geometry properties. Structure identification in terms of non-local characterization is done via area-based probability density functions of those geometrical properties. Classification is based on this geometrical characterization of individual structures and is enhanced via clustering techniques. Clustering algorithms allow the eduction of groups of structures without the need for strong a priori assumptions about their properties.

1.4 Choice of applications: passive scalar, enstrophy, and dissipation fields

Presently we apply this methodology, first, to a passive scalar advected and diffused in statistically stationary homogeneous isotropic turbulence with a mean scalar gradient imposed. Second, we study the structures of the enstrophy and dissipation fields obtained from homogeneous isotropic turbulence decaying in time. In all cases the flow is incompressible. The databases under analysis were obtained by DNS.

The dynamics of the passive scalar, c, are governed by the linear advection-diffusion equation:

$$\frac{\partial}{\partial t}c(\boldsymbol{x},t) + u_j(\boldsymbol{x},t)\frac{\partial}{\partial x_j}c(\boldsymbol{x},t) = D\frac{\partial^2}{\partial x_j \partial x_j}c(\boldsymbol{x},t), \tag{1.1}$$

where $\{u_j, j = 1,2,3\}$ are the components of the velocity field, $\boldsymbol{u}$, $\boldsymbol{x}$ is the position vector ($\{x_j, j = 1,2,3\}$ are the spatial coordinates), t is the time variable, and D is the diffusivity. In the presence of a uniform mean scalar gradient of magnitude μ_c in the x_1 direction—which will be preserved by the flow (see Corrsin, 1952)—the passive scalar can be split into its mean component, $\mu_c x_1$, and the passive scalar fluctuation, $c'(\boldsymbol{x},t)$. Thus $c(\boldsymbol{x},t) = \mu_c x_1 + c'(\boldsymbol{x},t)$ and the passive scalar fluctuation is then governed by:

$$\frac{\partial}{\partial t}c'(\boldsymbol{x},t) + u_j(\boldsymbol{x},t)\frac{\partial}{\partial x_j}c'(\boldsymbol{x},t) = D\frac{\partial^2}{\partial x_j \partial x_j}c'(\boldsymbol{x},t) - \mu_c u_1(\boldsymbol{x},t). \tag{1.2}$$

The mean scalar gradient acts as a source term for the scalar fluctuation, and a statistically stationary state can be reached (see Overholt & Pope, 1996). Passive scalars are of paramount importance in turbulent mixing and combustion and a vast effort has been dedicated to their study (see Warhaft, 2000, and the references therein). We choose it as a first case of application of our methodology for being a scalar field itself, governed by a relatively simple equation, before moving to other scalar fields derived from the velocity gradient tensor, with more complicated dynamics, such as the enstrophy and dissipation.

The analysis of the enstrophy and kinetic energy dissipation fields has been recurrent in the study turbulence through experiments (e.g., Zeff et al., 2003), numerical simulations (e.g., Ishihara et al., 2003), and theoretical developments (e.g., Pullin & Saffman, 1997; He et al., 1998; Wu et al., 1999). They are obtained, up to scaling factors, from the double contraction of the rotation- and strain-rate tensor fields. Physically, enstrophy and dissipation correspond to the remaining Galilean-invariant degrees of freedom of fluid particles, rotation and strain, once the dilatation is restricted for incompressible flows. This separation is useful but it does not decouple the equations of fluid motion. On the contrary, both fields appear highly coupled in the equations describing the dynamics of each

other (see Appendix A). In fact, the interaction between strain and rotation is intrinsic to the very nature of three-dimensional turbulence; in particular, vortex-stretching occurs when the strain-rate field stretches and amplifies vorticity. A study and comparison of the geometry of structures of both fields, at different scales, might be valuable in our understanding of turbulence. For that reason, we choose them as the second case of application of our methodology.

1.5 Grid resolution effects

Because of its multi-scale nature, a complete study of turbulence requires, both in experiments and in numerical simulations, spatial resolution that resolves the flow up to dissipation scales. A traditional grid resolution criterion used in DNS of homogeneous turbulence in a periodic box, for example, consists in resolving the flow up to scales of the order of the (average) Kolmogorov length scale. But in addition to being multi-scale, turbulence also shows intermittency (Batchelor & Townsend, 1949; Landau & Lifshitz, 1959; Kolmogorov, 1962): fluctuations of flow quantities can reach extreme amplitudes in short intervals of time and spatial distances. Furthermore, fluctuations of different amplitudes tend to cluster. Intermittency increases for higher Reynolds numbers (Okamoto et al., 2007) and also for smaller scales (Brasseur & Wang, 1992).

This suggests that the traditional grid resolution criterion, based on an average dissipation scale, might be inappropriate, since much smaller scales are locally present due to those high fluctuations. Therefore, the resolution required to resolve all scales of turbulent flows increases significantly (see Sreenivasan, 2004). This condition may become even more restrictive when studying the geometry of structures in turbulence, and is explored during the application of our methodology to the enstrophy and dissipation fields by means of databases corresponding to multiple numerical simulations performed at different resolutions but otherwise identical.

1.6 Structure interaction

The dynamics of sheets and tubes are greatly affected by their own interactions. Common examples are the coalescence and reconnection of approaching vortex tubes and the roll-up of vortex sheets to form vortex tubes resulting from the Kelvin-Helmholtz instability. These interactions among sheets and tubes can be seen as the translation of the strain-rotation interaction itself to the structural level of turbulence, and help explain the presence of intermittency and the process of multi-scale energy cascade in turbulence (see Kraichnan, 1974).

An interesting example of the geometrical relations between rotation- and strain-rate fields is the local alignment of the vorticity with the intermediate strain-rate eigendirection, for incompressible homogeneous isotropic turbulence. It was observed first in numerical simulations (Ashurst et al., 1987) and confirmed experimentally (see Tsinober et al., 1992; Tao et al., 2000). Theoretical explanations combine local and non-local arguments (see Jiménez, 1992; Nomura & Post, 1998; Hamlington et al., 2008). But this prevailing alignment between vorticity and the intermediate eigendirection of the strain-rate tensor is observed to switch towards the direction associated with the most negative[1] (compressional) eigenvalue of the strain-rate tensor at the ends of tube-like structures (Nomura & Post, 1998), which is consistent with the compressive straining of the vorticity occurring in those regions.

Other geometrical analyses regarding the proximity of different types of structures, in relation to their shapes, could be useful in further explaining those interactions and also improve structure-based models of the fine scales of turbulence. For that purpose, we have developed a methodology for the study of such proximity issues, from a geometrical viewpoint, among structures of different fields and scales. It takes advantage of many of the features of the methodology for the study of the geometry of structures in turbulence also introduced here.

[1] For incompressible flow the trace of the strain-rate tensor is null, ensuring at least one positive and one negative eigenvalue of that tensor.

1.7 Outline

This thesis is organized as follows: Chapter 2 describes the three main steps of the methodology for the non-local multi-scale study of the geometry of structures—extraction, characterization, and classification—with emphasis on the conceptual basis and on some particular implementation details. In Chapter 3, we present a system test that validates the methodology applied to a virtual world of modeled structures. Chapter 4 shows results of its first application to extended passive scalar structures educed from a DNS database of incompressible homogeneous isotropic turbulence stationary in time. We apply, in Chapter 5, the same methodology to structures of the enstrophy and dissipation fields, comparing the results of both fields, from another DNS database of incompressible homogeneous isotropic turbulence decaying in time. This database includes three different grid resolutions, allowing us to study how this parameter affects the geometry of educed structures and the validity of the traditional grid resolution criterion in DNS from a geometrical standpoint. In Chapter 6, we combine our non-local methodology with two local criteria of identification of vortex tubes and sheets in turbulence (Horiuti & Takagi, 2005) that are based on scalar fields obtained from the velocity gradient tensor. An assessment of the geometries expected from those local criteria is done. Chapter 7 introduces a new methodology for the study of the proximity of multiple sets of structures, also in terms of geometry and based on non-local measures through area-coverage quantification. We apply this methodology to the pairs of two scalar fields used by the local identification criteria in Chapter 6 and also to the enstrophy and dissipation fields, considering the multi-scale decomposition performed in Chapter 5. Chapter 8 summarizes the conclusions of this work and comments on its possible future directions. The contents of Chapters 2, 3, and 4, along with the corresponding conclusions included in Chapter 8, will appear in Bermejo-Moreno & Pullin (2008).

We emphasize that the tools developed here—the multiresolution analysis, geometric characterization, spectral projection, clustering algorithms, and proximity analysis—can be applied to many scalar and tensor fields in turbulence, and in fields beyond fluid dynamics.

Chapter 2

Methodology for the study of the geometry of structures in turbulence

The starting point of this methodology is a three-dimensional scalar field obtained from a turbulence database. Three properties were sought in its development: multi-scale capability, non-local character, and geometry-based analysis. It consists of three main steps: extraction, characterization, and classification of structures. They are explained in each of the sections of this chapter.

2.1 Extraction of structures

The main requirement imposed on the extraction process is to enable eduction of structures associated with different ranges of scales. Although scale decomposition is commonly defined in Fourier space, the nature of Fourier basis functions, that are localized in wavenumber but not in physical space, makes top-hat window filtering in Fourier space inappropriate for the purpose of educing structures that are extended but compact in physical space. Thus, a transformation with basis functions that are localized both in Fourier space, where the ranges of scales are defined, and in physical space, where the structures are to be educed, is required. For this purpose, the curvelet transform (Candès & Donoho, 2003*a*,*b*) in its three-dimensional discretized version (Ying et al., 2005; Candès et al., 2005) is used. Owing to the multi-dimensional character of their definition, curvelets, unlike wavelets, are naturally suited for detecting, organizing, or providing a compact representation of

intermediate multi-dimensional structures.

2.1.1 The curvelet transform

Curvelets, the basis functions of the curvelet transform, are localized in scale (frequency/Fourier space), position (physical space) and orientation (unlike wavelets). The frequency space is smoothly windowed in radial and angular spherical coordinates, providing the decomposition in different scales and orientations, respectively. For a given scale[1], j, the radial window smoothly extracts the frequency contents near the dyadic corona $[2^{j-1}, 2^{j+1}]$. A low-pass radial window is introduced for the coarsest scale, j_0. The unit sphere representing all directions in $\mathbb{R}^3$ is partitioned, for each scale $j > j_0$, into $O\left(2^{j/2} \cdot 2^{j/2}\right) = O\left(2^j\right)$ smooth angular windows, each with a disk-like support of radius $O\left(2^{-j/2}\right)$. In a discrete three-dimensional data field, of uniform grid of size n^3, the last scale, j_e, which extracts the highest frequency content, is given by $j_e = \log_2(n/2)$.

Denoting by $f(n_1, n_2, n_3)$ the scalar field, where $0 \leq n_i < n$, being n the number of grid points in each direction, the discrete version of the curvelet transform (see Ying et al., 2005) provides a set of coefficients $c^D(j, l, k)$ defined as

$$c^D(j, \ell, k) \equiv \sum_{n_1, n_2, n_3} f(n_1, n_2, n_3)\, \overline{\varphi^D_{j,\ell,k}(n_1, n_2, n_3)} \tag{2.1}$$

where $j, \ell \in \mathbb{Z}$, $k = (k_1, k_2, k_3)$ (j represents the scale, ℓ the orientation, and k the spatial location); $\varphi^D_{j,\ell,k}(n_1, n_2, n_3)$ are the curvelets, defined in Fourier space by

$$\hat{\varphi}^D_{j,\ell,k}(\omega) \equiv \tilde{U}_{j,\ell}(\omega)\, \exp\left(\frac{-2\pi i \sum_{i=1}^{3} \frac{k_i \omega_i}{L_{i,j,\ell}}}{\sqrt{\prod_{i=1}^{3} L_{i,j,\ell}}} \right) \tag{2.2}$$

for $\{0 \leq k_i < L_{i,j,\ell}, i = 1, 2, 3\}$—where ω is the wavenumber; $\tilde{U}_{j,\ell}(\omega)$ is the frequency window $\tilde{U}_{j,\ell}(\omega) = \tilde{W}_j(\omega)\, \tilde{V}_{j,\ell}(\omega)$, being $\tilde{W}_j(\omega)$ and $\tilde{V}_{j,\ell}(\omega)$ the radial and angular windows; and $\{L_{i,j,\ell}, i =$

[1] The term scale, when referred to the index j in curvelet space, denotes in fact the range of scales in physical space that results from the radial window filter in Fourier space.

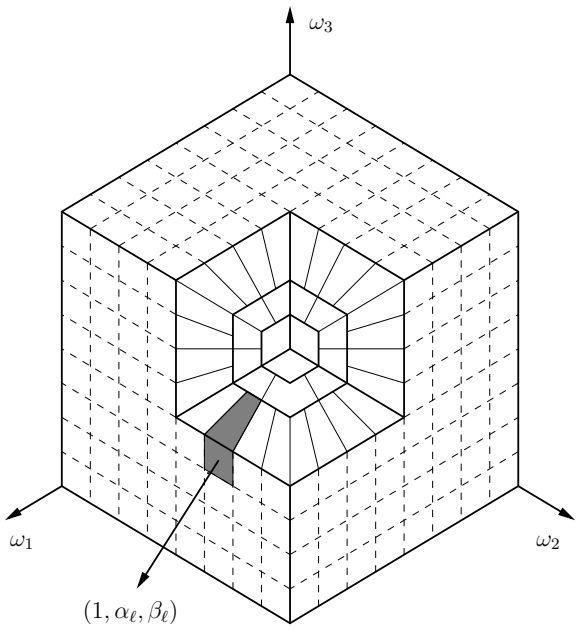


Figure 2.1: Frequency window $\tilde{U}_{j,\ell}$ (darkened region) defined in the three-dimensional discrete curvelet transform, extracting the frequency content near the wedge with center slope $(1, \alpha_\ell, \beta_\ell)$ (figure based on Ying et al. (2005))

$1, 2, 3\}$ are three positive integers such that: $i)\Big\{ \nexists\, \omega, \omega' \;\Big|\; \tilde{U}_{j,\ell}(\omega) \geq \tilde{U}_{j,\ell}(\omega') \text{ and } \omega_i \text{ is multiple of } L_i, i = 1, 2, 3 \Big\}$;

$ii)\Pi_{i=1}^{3} L_{i,j,\ell}$ is minimal. A Cartesian coronae is used, so that:

$$\tilde{W}_{j_0}(\omega) = \Phi_{j_0}(\omega); \qquad \tilde{W}_j(\omega) = \sqrt{\Phi_{j+1}^2(\omega) - \Phi_j^2(\omega)}, \qquad j > j_0, \tag{2.3}$$

where $\Phi_j(\omega_1, \omega_2, \omega_3) = \phi(2^{-j}\omega_1)\, \phi(2^{-j}\omega_2)\, \phi(2^{-j}\omega_3)$, and ϕ is a smooth function such that $0 \leq \phi \leq 1$: it equals unity on $[-1, 1]$ and zero outside $[-2, 2]$. The angular window for the ℓth direction is defined (for example, in the $\omega_1 > 0$ face of the unit cube) as

$$\tilde{V}_{j,\ell}(\omega) = \tilde{V}\left(2^{j/2}\frac{\omega_2 - \alpha_\ell \omega_1}{\omega_1}\right) \tilde{V}\left(2^{j/2}\frac{\omega_3 - \beta_\ell \omega_1}{\omega_1}\right) \tag{2.4}$$

where $(1, \alpha_\ell, \beta_\ell)$ is the direction of the center line of the wedge (see Figure 2.1) defining the center slope for the ℓth wedge. Wherever three smooth angular windows $\tilde{V}_{j,\ell}$, $\tilde{V}_{j,\ell'}$, and $\tilde{V}_{j,\ell''}$ overlap, they are redefined as $\left(\tilde{V}_{j,\ell}, \tilde{V}_{j,\ell'}, \tilde{V}_{j,\ell''}\right) / \sqrt{\tilde{V}_{j,\ell}^2 + \tilde{V}_{j,\ell'}^2 + \tilde{V}_{j,\ell''}^2}$.

$\tilde{W}_j(\omega)$ and $\tilde{V}(\omega)$ satisfy:

$$\sum_{j \geq j_0} \tilde{W}_j^2(\omega) = 1, \qquad \sum_{\ell=-\infty}^{\infty} \tilde{V}^2(t - 2\ell) = 1. \tag{2.5}$$

Curvelets form a tight-frame in $L^2\left(\mathbb{R}^3\right)$. Any function $f \in L^2\left(\mathbb{R}^3\right)$ can be expanded as $f = \sum_{j,\ell,k} \langle \varphi_{j,\ell,k}, f \rangle \varphi_{j,\ell,k}$, where $\varphi_{j,\ell,k}$ is the curvelet at scale j, orientation ℓ, and position $k = (k_1, k_2, k_3)$. Parseval's identity holds: $\sum_{j,\ell,k} \mid \langle f, \varphi_{j,\ell,k} \rangle \mid^2 = \|f\|^2_{L^2(\mathbb{R}^3)}$. The effective longitudinal and cross-sectional dimensions (length and width), of curvelet basis functions in physical space follow the relation width$\approx$length^2 (parabolic scaling). As a consequence of this parabolic scaling, curvelets are an optimal (sparse) basis for representing surface-like singularities of codimension one. These are three of the most remarkable properties of the curvelet transform.

We apply the curvelet transform to a scalar field, obtained from a turbulence database at an instant in time, but again emphasize its broader applicability to other fields. The curvelet transform allows a multi-scale decomposition by filtering in curvelet space the different scales of interest $j = j_0, ..., j_e$, individually or in groups. In addition, for anisotropic fields with privileged direction(s) (e.g., shear flows), a multi-orientation decomposition may be useful for studying structures according to their directionality (by using the angular window filtering in frequency space of the curvelet transform (index ℓ in curvelet space)). Throughout this thesis, only the multi-scale decomposition is used, which could be also attained by other multi-resolution techniques sharing the same choice of sub-band radial filtering decomposition in Fourier space. Nevertheless, those capabilities that set curvelets apart from other multi-resolution techniques, e.g., multi-orientation decomposition and compact representation of surface-like singularities, justify its early implementation within the frame of this methodology, enhancing its potential applications and possibilities of expansion.

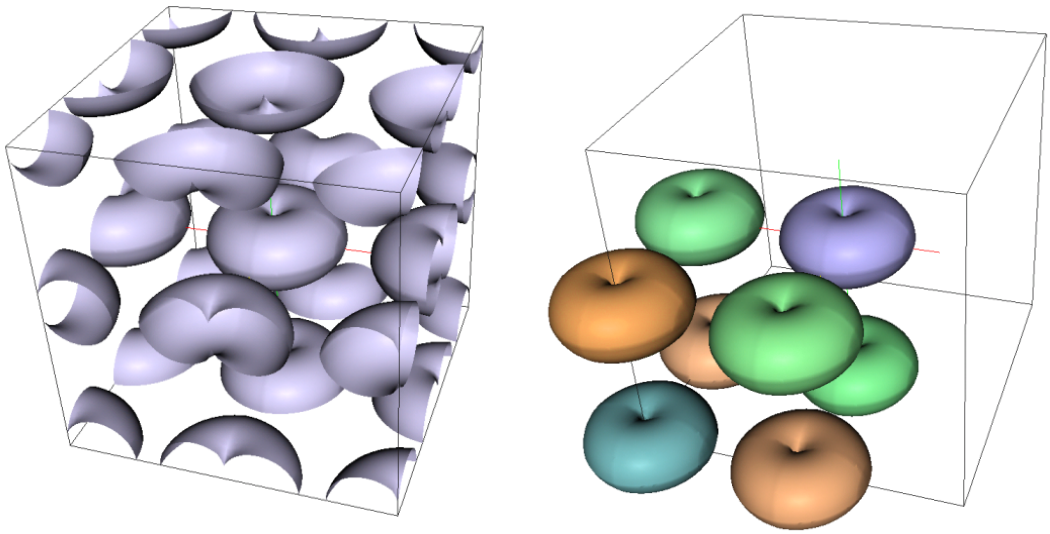
For each scale $j = j_0, ..., j_e$, a new scalar field is obtained after filtering all other scales ($j\prime \neq j$) in curvelet space and inverse transforming to physical space. Thus, a set of $j_e - j_0 + 1$ filtered scalar fields results from the original field. The volume-based probability density functions (pdfs) of the filtered fields are, in general, different from each other and from the original field; their comparison

can be useful in determining how the original scalar field is distributed among the different scales.

After this multi-scale analysis, a second step is applied in the extraction process, by which the structures of interest associated with each relevant range of scales are educed. Currently those structures of interest are defined as the individual disconnected surfaces obtained by iso-contouring each filtered scalar field at particular contour values (for example, the mean value of that filtered scalar field plus a multiple of its standard deviation). See Appendix B for a physical interpretation of the educed structures following this multi-scale decomposition plus iso-contouring procedure.

2.1.2 Periodic reconnection

In the case of scalar fields with periodic boundaries, an additional step is included in the extraction process, to reconnect those structures intersecting boundaries with their periodic continuation on the opposite boundaries. Figure 2.2 shows an example of the application of such periodic reconnection algorithm to a set of 3D structures obtained from a periodic scalar field.



(a) (b)

Figure 2.2: Example of application of the periodic reconnection algorithm to a set of boundary-intersecting structures obtained from a periodic three-dimensional scalar field. (a) Before periodic reconnection. (b) After periodic reconnection, where the color of each structure indicates the number of pieces involved in the reconnected structure for this particular scenario: blue = 1 (non-intersecting), green = 2, orange = 4, cyan = 8

Structures spanning across multiple repetitions of the periodic domain are accounted for in the

algorithm, as shown schematically in Figure 2.3 for a 2D example. The only case that cannot be completely reconnected is the structure with infinite extent.

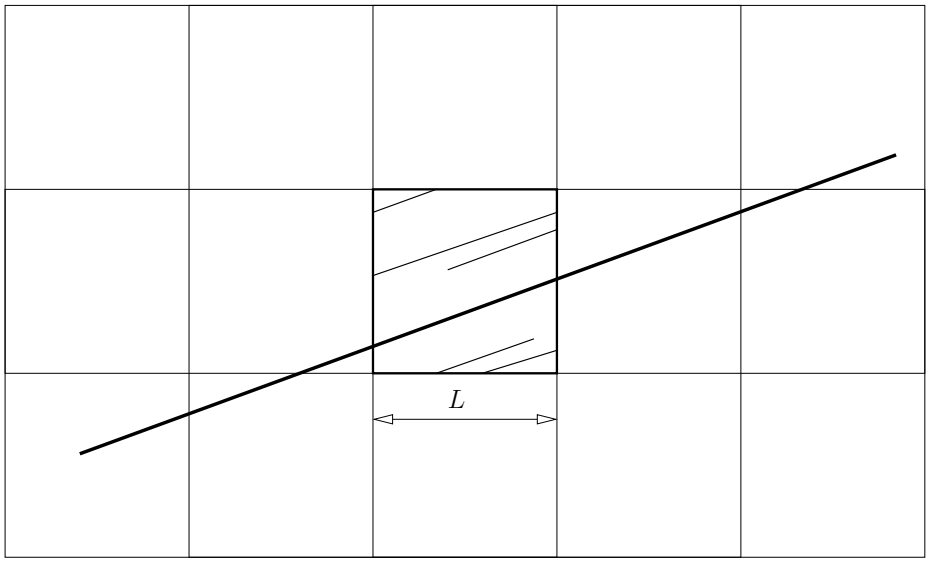


Figure 2.3: 2D example of reconnection of a periodic structure spanning across multiple extensions of the periodic domain. The original fragments of the structure are represented in the original domain (central square). The resulting structure, after reconnection, is represented by the thick line.

2.2 Characterization of structures

We seek a geometrical characterization able to distinguish structures based on their shape. A two-step method is used: first, a suitable set of differential-geometry properties is obtained locally (at all points of the surface), and then area-based probability density functions of those local properties are calculated, making the transition from local to non-local (in the surface sense) possible.

2.2.1 Shape index and curvedness

Shape index, Υ, and curvedness, Λ, (see Koenderink & van Doorn, 1992) are the differential-geometry properties chosen to represent locally the geometry of the surface. They are related to the *principal curvatures* $\{\kappa_1, \kappa_2\}$ of a surface at a given point by:

$$\Upsilon \equiv -\frac{2}{\pi} \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right), \qquad \Lambda \equiv \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}. \tag{2.6}$$

Υ is scale-independent, whereas Λ is scale-dependent, having the dimensions of a reciprocal length. The scaling is such that, for example, Λ at every point on a sphere equals the absolute value of its reciprocal radius, $1/R$, whereas the cylinder of radius R presents $\Lambda = 1/(\sqrt{2}R)$ for all points. The principal curvatures, $\{\kappa_1, \kappa_2\}$, are obtained as the maximum and minimum values of the normal curvature, κ_n, in all possible directions of the tangent plane defined at the point P of the surface of study. The normal curvature, κ_n, at a point P in a given direction $\boldsymbol{a}$ of the tangent plane, defined as the division of the second and first fundamental forms of differential geometry applied in that direction, $\boldsymbol{a}$, can also be interpreted as the inverse of the radius of curvature, R, of the curve obtained as the intersection of the surface and the plane defined by the direction $\boldsymbol{a}$ and the normal $\boldsymbol{N}$ to the surface at the point P. Thus higher values of the curvedness correspond to smaller radius of curvature (and, therefore, more locally curved surface at P). All regular patches of a regular surface M map on the domain $(\Upsilon, \Lambda) \in [-1, +1] \times \mathbb{R}^+$, except for the planar patch, which has an indeterminate shape index and nil curvedness (since $\kappa_1 = \kappa_2 = 0$).

The mapping $(\kappa_1, \kappa_2) \rightarrow (\Upsilon, \Lambda)$ represents (see Figure 2.4) a transformation from Cartesian coordinates (κ_1, κ_2) to non-standard polar coordinates (Υ, Λ). For any point in the (κ_1, κ_2) plane, Υ contains the information on the direction (measured as the angle, ϕ, with respect to the axis $\kappa_1 - \kappa_2$, rescaled into the range $[-1, +1]$ by $\Upsilon = -2\phi/\pi$), whereas Λ contains the information on the distance, ϱ, to the origin (rescaled as $\Lambda = \varrho/\sqrt{2}$). The convention chosen when ordering the principal curvatures ($\kappa_1 \geq \kappa_2$) implies that only the region $\kappa_1 - \kappa_2 \geq 0$ of the (κ_1, κ_2) plane is accessible (see Figure 2.4). Therefore, the polar angle ϕ can only have values in the range $[-\pi/2, +\pi/2]$, and, consequently, $\Upsilon = -\frac{\phi}{\pi/2} \in [-1, +1]$ covers all the possible cases, thus making the mapping $(\kappa_1, \kappa_2) \rightarrow (\Upsilon, \Lambda)$ injective (excluding the point $(\kappa_1, \kappa_2) = (0, 0)$ from its domain) by eliminating the multi-valuedness of the arctan function used in the definition of the shape index. The absolute value of the shape index $S \equiv |\Upsilon|$ represents the local shape of the surface at the point P, with $0 \leq S \leq 1$. Its sign indicates the direction of the normal, distinguishing, for example, convex from concave elliptical points. Figure 2.5 shows the range of values of Υ and sketches of the local shapes associated with representative values, with the names of corresponding points. Figure 2.6 shows

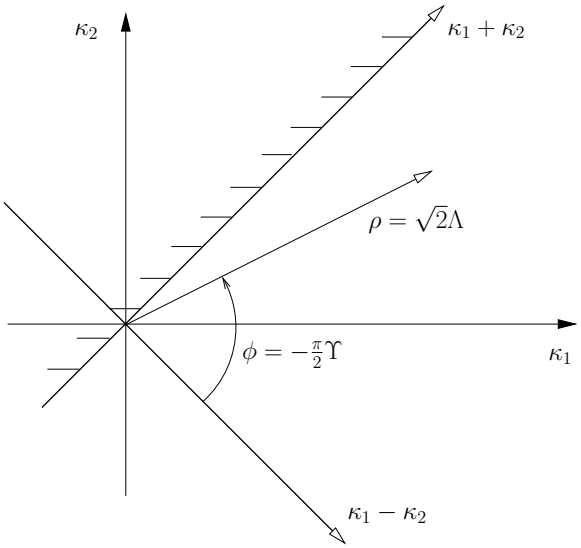


Figure 2.4: Transformation from (κ_1, κ_2) to (Υ, Λ)

the mapping of both Υ and Λ in the plane of principal curvatures, also with representative local shapes. A deeper mathematical background of these differential-geometry concepts is presented in Appendix C.

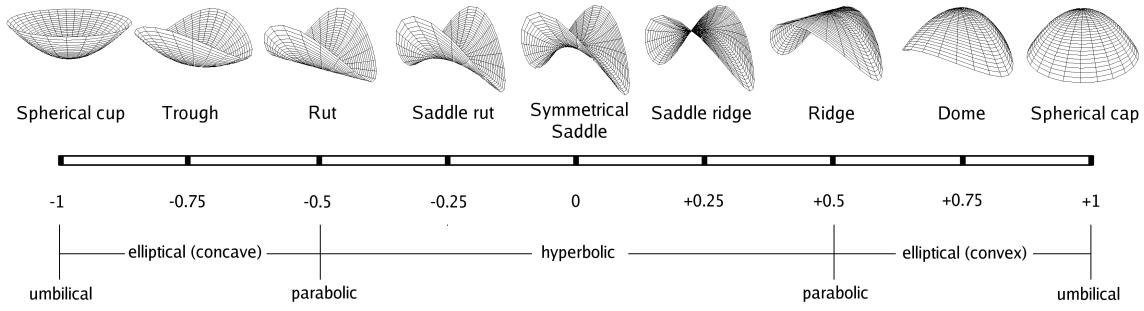


Figure 2.5: Range of shape index, Υ, with its most representative associated local shapes (figure based on Koenderink & van Doorn (1992))

2.2.2 Joint probability density function (jpdf)

From the pointwise Υ and Λ, a two-dimensional area-based joint probability density function in the space of (S, Λ) can be obtained (see Appendix D). Since Λ is scale dependent, in order to compare the shape of surfaces of different sizes, a non-dimensionalization is required for each surface. Selection of the appropriate length scale for this purpose is critical; several can be obtained from global

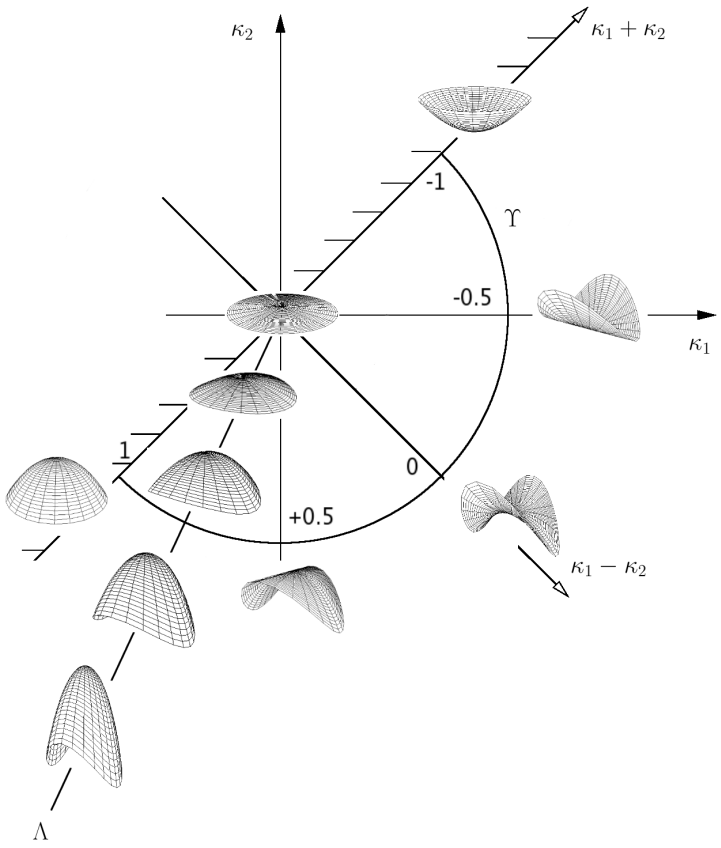


Figure 2.6: Representative local shape in the combined (κ_1,κ_2) plane

geometrical invariants of the surface, such as the square root of its area (A), the cubic root of its volume (V), etc. Presently we define

$$C \equiv \mu\Lambda, \qquad \mu \equiv 3\,\frac{V}{A}. \tag{2.7}$$

For the sphere, $C = 1$. The definition of a volume implies that the structure under consideration is a closed surface. Thus, only closed surfaces educed from the scalar field are studied. For periodic domains only those structures with infinite extent will not be closed. All others, following periodic reconnection, will be closed. For non-periodic domains of limited extent, those structures intersecting boundaries will not be closed, but they could still be considered in the analysis by closing them either with the boundaries that they intersect or with their mirrored extension across those boundaries,

for example.

Another dimensionless global parameter useful in the characterization of the geometry of the educed closed structures is

$$\lambda \equiv \sqrt[3]{36\pi}\,\frac{V^{2/3}}{A}. \tag{2.8}$$

It represents the stretching of the structure; the lower its value, the more stretched the structure is. For the sphere $\lambda = 1$.

The area-based jpdf $\mathcal{P}(S,C)$, $\int\int \mathcal{P}(S,C)\,\mathrm{d}S\mathrm{d}C = 1$, contains non-local information on the geometry of the surface. $\mathcal{P}(S,C)$ can be geometrically interpreted as a representation of how the local shape, S, is distributed across the different (relative) length scales present on the surface, given by C, in terms of area coverage. For closed surfaces, their geometry and topology are related by the Gauss–Bonnet theorem, which imposes an integral constraint on the area-based joint probability density function of S and C (see Appendix E for details).

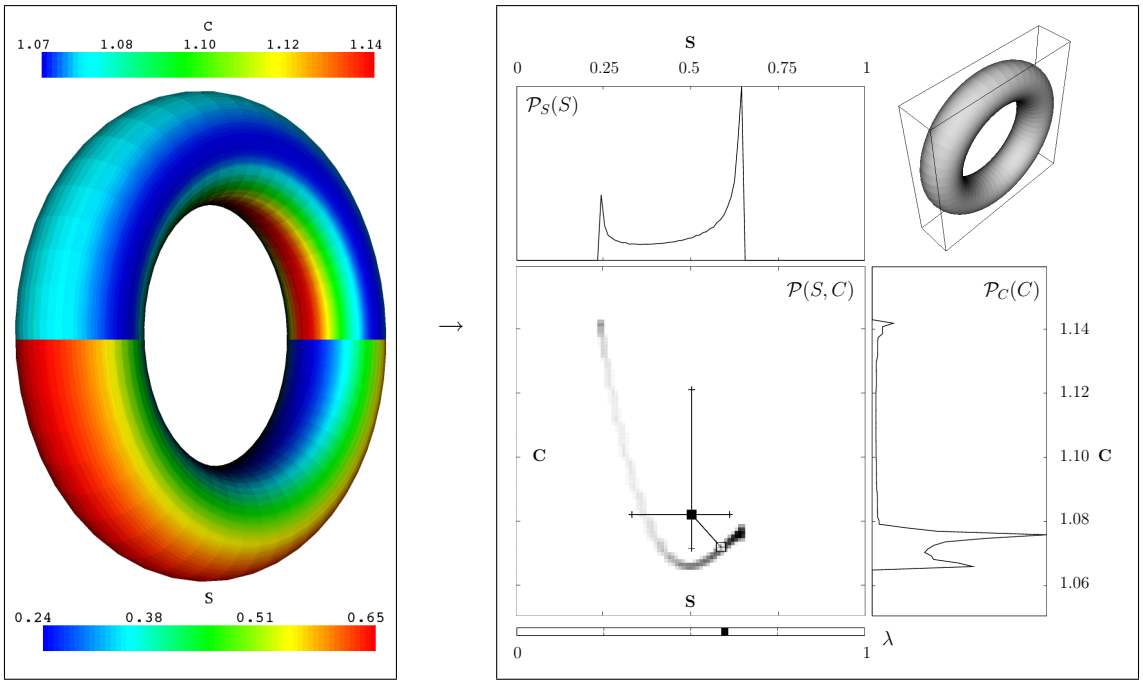
2.2.3 Signature of a structure

We consider $\mathcal{P}(S,C)$ plus its associated one-dimensional marginal pdfs,

$$\mathcal{P}_{\mathcal{S}}(S) = \int \mathcal{P}(S,C)\,\mathrm{d}C, \qquad \mathcal{P}_{\mathcal{C}}(C) = \int \mathcal{P}(S,C)\,\mathrm{d}S, \tag{2.9}$$

to be the signature of the structure. This is complemented with its area A and λ, representing the stretching of the structure. We find it useful to display $\mathcal{P}(S,C)$ mapped onto the (S,C)-plane with greyscale rendering of $\mathcal{P}$; white $\equiv$ 0, black $\equiv \max(\mathcal{P})$. Additionally, we plot $\mathcal{P}_S(S)$ and $\mathcal{P}_C(C)$ on the S (top) and C (right) axes respectively; see Figure 2.7 for an example. This geometrical characterization is based on properties of the structure that are invariant with respect to translations and rotations of the reference system, and therefore, are suited for comparing structures based on their geometry, the basis of the next step of this methodology: the classification of structures.

Several methods have been proposed in the computer graphics literature for estimating curvatures of a discretized surface (such as the ones that represent our structures in the computational domain).



(a) (b)

Figure 2.7: Example of a three-dimensional surface (a) with S and C mapped onto it (bottom and upper halves, respectively) and its corresponding signature (b), for which a projection of the three-dimensional physical structure is shown at the top-right corner, its area-based joint two-dimensional probability density function (in terms of S and C) is presented in the bottom-left area, while the marginal probability density functions of both S and C, are drawn at its top and right sides, respectively. The value of the stretching parameter, λ, is represented below the jpdf by a black bar (in a scale from 0 to 1). Mean and feature centers, as well as upper and lower distances for each variable of the jpdf, are superimposed to the jpdf, as the filled and hollow squares, respectively (refer to §2.3)

A subset of them, applied to the case in which the discretized surface is a triangular mesh, was here implemented and tested (Chen & Schmitt, 1992; Dong & Wang, 2005; Taubin, 1995; Meyer et al., 2003). Finally, a modification of the algorithm proposed by Dong & Wang (2005) (based on Chen & Schmitt (1992)) is used. The only modification is the way in which the normal vector to each face of the discretized surface is computed, following the method proposed by Chen & Wu (2004).

2.3 Classification of structures

A process of classification assigns different elements of a given set to groups based on the similarities of their signatures. In our system, the elements to be classified are the educed structures, and the

signatures are given by $\mathcal{P}(S,C), \mathcal{P}_S(S), \mathcal{P}_C(C)$, and λ, obtained in the characterization step for each structure. Among the different approaches to the problem of classification, we seek those involving as little a priori knowledge as possible of the relationships governing the different groups and of the number of groups present in the set of elements under evaluation. This leads to the utilization of learning-based clustering techniques. The idea behind this approach is to be able to detect other types of geometries apart from the known tube-like and sheet-like structures in turbulence databases, should they exist, by not imposing strong assumptions on the groups.

2.3.1 Clustering algorithm

The clustering algorithm used in this classification step combines several techniques found in the data mining, pattern recognition, and artificial intelligence literature (see, for example, Berkhin, 2002, for a survey of such clustering techniques). It is a locally scaled spectral partitional clustering algorithm that automatically determines the number of clusters. Its main steps are summarized below using the notation proposed by Ng et al. (2001) in their NJW algorithm, that conforms the core of our technique. Additions, particularizations, and modifications to the NJW algorithm are also described below. In what follows we denote a set of N structures at a particular scale by the N elements $E = \{e_1, \ldots, e_N\}$. For each member of this set we construct a set of parameters $\{p[k], k = 1, \ldots, N_P\}$ which will serve as the contracted computational signature of the structure. These will be a finite set of moments of $\mathcal{P}(S,C), \mathcal{P}_S(S), \mathcal{P}_C(C)$, to be defined subsequently, together with λ. The $p[k]$ will also define a *feature space of parameters* in which the elements e_i are mapped. Typically $N = O(10^2 - 10^5)$, depending on the scale, and it will be seen that $N_p = 7$.

1. Start from a set of N elements $E = \{e_1, \ldots, e_N\}$ and their corresponding contracted signatures $\{p_{e_i}[k], k = 1, \ldots, N_p\}$.

2. Construct the *distance matrix*, $d_{ij} = d(e_i, e_j)$, $e_i, e_j \in E$. The element d_{ij} of the distance matrix measures dissimilarity between the two elements e_i and e_j of E, based on their signatures.

Presently we define the distance

$$d_{ij} = F(\{p_{e_i}[k] - p_{e_j}[k], k = 1, \ldots, N_p\}) \tag{2.10}$$

where $p_{e_l}[k]$ is the kth parameter associated with element e_l. The weighting function F defines a distance in that space of parameters. For example, a functional dependence of F of the form $F(\boldsymbol{x}) = \left(\sum_i x_i^2\right)^{1/2}$ defines a Euclidean distance in the feature space of parameters.

3. Construct a *locally scaled affinity matrix* $\hat{\boldsymbol{A}} \in \mathbb{R}^{N \times N}$ defined by

$$\hat{A}_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma_i \sigma_j}\right) \tag{2.11}$$

where σ_l is a *local scaling parameter* introduced by Zelnik-Manor & Perona (2005) and defined as the distance of the element e_i to its rth closest neighbor, denoted by $e_{r,i}$, $\sigma_i = d(e_i, e_{r,i})$ (a value of $r = 7$ is used, following Zelnik-Manor & Perona, 2005). The purpose of introducing a local scaling parameter is to take into consideration the multiple scales that can occur in the clustering process, which is important, for example, when tight clusters are embedded within more sparse background clusters. Note that the elements of the diagonal of $\hat{\boldsymbol{A}}$ are null.

4. Normalize $\hat{\boldsymbol{A}}$ with a diagonal matrix $\boldsymbol{D}$ such that $D_{ii} = \sum_{j=1}^{N} \hat{A}_{ij}$, obtaining the *normalized locally scaled affinity matrix* $\boldsymbol{L} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$

5. For N_C varying between the minimum and maximum number of clusters considered, do the following loop:

   (i) Find the N_C largest eigenvectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_C}\}$ of $\boldsymbol{L}$ and form the matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_C}] \in \mathbb{R}^{N \times N_C}$. This step constitutes the spectral part of the algorithm. It is intended to map the elements e_i onto a different eigenspace where clusters can be better identified. It can be considered a pre-clustering step that, combined with the local scaling explained in one of the previous steps, allows clustering of elements with more complicated relationships among them (and to other clusters) than traditional clustering techniques that do not

use these features. For example, concentric clusters can be easily educed by means of spectral clustering.

(ii) Re-normalize the rows of $\boldsymbol{X}$ so that they have unitary length, obtaining the matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times N_C}$ as $Y_{ij} = X_{ij} / \left(\sum_j X_{ij}^2 \right)^{1/2}$.

(iii) Treat each row of $\boldsymbol{Y}$ as a point in $\mathbb{R}^{N_C}$ and cluster them into N_C clusters via K-means algorithm.

(iv) Assign the original element e_i to cluster k if row i of $\boldsymbol{Y}$ was assigned to cluster k in the previous step.

(v) Obtain *optimality score* for this number of clusters N_C (see §2.3.3).

6. After the previous step has been done for all the possible numbers of clusters under evaluation, determine the optimum number of clusters based on the minimization of the optimality score for each one of the possible numbers of clusters (as will be described in §2.3.3).

The K-means clustering algorithm mentioned above is one of the simplest partitional clustering techniques available. It first initializes the cluster centers (for a given number of them). Then it assigns each element to the cluster with the closest centroid to that element. After all elements have been assigned, it recalculates the position of the cluster centers. The last two steps are repeated until the cluster centers no longer move. Different implementations of the K-means clustering algorithm differ mainly in the initialization of the cluster centers: we choose the initial position of the first cluster center randomly among all the elements; initial positions of subsequent cluster centers are obtained as the farthest elements to the previously assigned cluster centers. Several initializations following that procedure are performed to avoid local minima.

2.3.2 Feature and visualization spaces; definition of the $p[k], k = 1, ..., N_p$

The selection of the $p[k]$ used to define the feature space plays a decisive role in the classification step. Each structure will be represented by a point in that feature space and its distance to the other points will define the similarity to their corresponding structures. The number of parameters

(dimensions of feature space) should be sufficiently large to distinguish satisfactorily relevant groups of structures, but at the same time, it should be kept as small as possible to avoid the so-called 'curse of dimensionality' (see Bellman, 1961) that affects unsupervised learning algorithms, like the clustering method used in this methodology, compromising its success by making the points too disperse in such high-dimensional space. The set of (seven) parameters chosen here for each element e_i of E is

$$\{p[k], k = 1, \ldots, 7\} \equiv \{\hat{S}, \hat{C}, \lambda, d_u^S, d_l^S, d_u^C, d_l^C\} \tag{2.12}$$

where $\hat{S}, \hat{C}$ denotes the *feature center* of $\mathcal{P}(S, C)$ and d_u^S, d_l^S, d_u^C, d_l^C are the *upper* and *lower distances* of the jpdf in each variable. The feature center ($\hat{S}$,$\hat{C}$) takes into account the asymmetry of the jpdf, correcting the mean center ($\bar{S}, \bar{C}$) so that the feature center lies closer to the region of higher density of the jpdf. The upper and lower distances, d_u and d_l, can be regarded as the r.m.s. of the part of the jpdf above and below, respectively, its mean value. A graphical example can be seen in Figure 2.7, where the mean and feature centers have been superimposed to their corresponding jpdf. Definitions of feature center, upper and lower distances, together with a representative one-dimensional example can be found in Appendix F.

Based on the idea of the feature space of parameters used for educing clusters of similar structures, we define a *visualization space*, intended to provide a graphical representation of the distribution of individual structures in a three-dimensional space, providing qualitative and quantitative information. In general, the higher-dimensional character of the feature space prevents its use as visualization space, but the utilization of *glyphs*, scaling, and coloring allows more than just three dimensions to be represented in the visualization space.

We define the three axes of the visualization space by $\hat{S}$, $\hat{C}$, and λ. Owing to the choice of non-dimensionalization of the curvedness and the normalization factors (see §2.2), as well as the intrinsic meaning of the shape index, curvedness, and stretching parameter, it is possible to identify regions in the visualization space with a particular geometrical meaning for those structures whose representation lies in them. For example, blob-like structures occupy the region near the point

$(1, 1, 1)$ (which corresponds to spheres); tube-like structures are localized near the $(1/2, 1, \lambda)$[2] axis (λ being an indication of how stretched the tube is) and the transition to sheet-like structures occurs as the curvedness and λ decrease. The plane $\hat{C} = 0$ is the limiting case of planar structure; furthermore, any structure composed of (predominantly) planar regions, thus featureless in the curvature sense, will have a (nearly) nil $\hat{C}$, independently of its relative aspect ratios, that will nevertheless affect its λ value. See Appendix G for an analysis of these limiting values. Throughout this thesis, the visualization space is presented by a set of two-dimensional projections (see Figure 2.8 for an example).

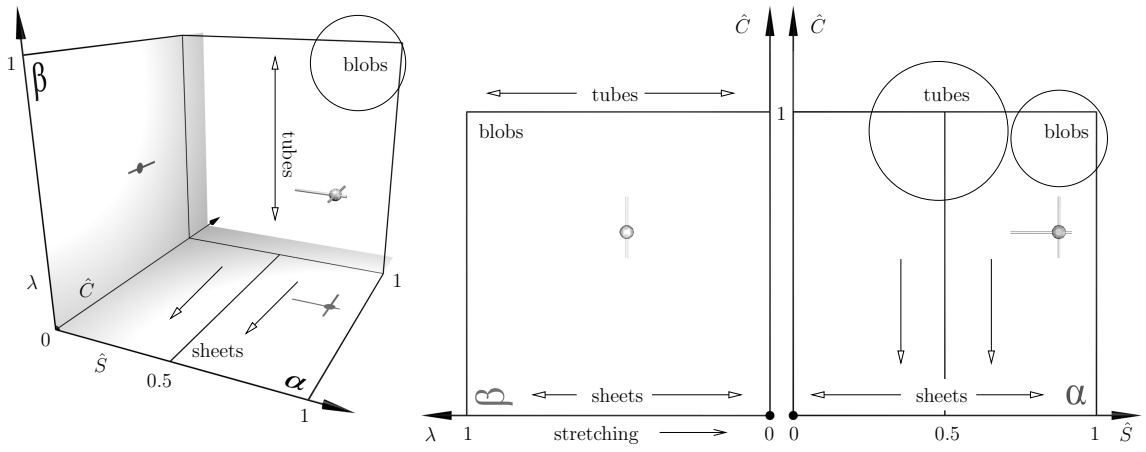


Figure 2.8: Projections of the *visualization space* with the predominantly blob-, tube- and sheet-like regions sketched: three-dimensional perspective projection (left), two-dimensional orthogonal projections (right) of the planes β (formed by the axes $\hat{C}$ and λ) and α (formed by $\hat{S}$ and $\hat{C}$). For example, a glyph consisting of a sphere and four bars along the $\pm\hat{S}$, $\pm\hat{C}$ axes can represent nine parameters of the characterization of the corresponding structure: $\hat{S}, \hat{C}, \lambda$ given by the center of the sphere, upper and lower distances of S and C given by each bar, the surface area A of the associated structure, given by the size of the glyph, and the group to which the structure belongs, given by the color of the glyph

2.3.3 Optimality score: silhouette coefficient

The determination of the optimum number of clusters is based on the minimization of an *optimality score*. Different approaches have been considered. Among them, probabilistic criteria that consider the relative increment of complexity of a model (set of clusters) when another parameter (cluster)

[2]Note that a value of the shape index equal to 1/2 corresponds to locally cylindrical shapes, that are predominant in tube-like structures. The dimensionless curvedness of a straight elongated circular cylinder of radius R reduces to $C \simeq 3V/A\sqrt{2}R \approx 3/2\sqrt{2} \approx 1.06$.

is added, such as the *Bayesian Information Criterion (BIC)* BIC (Schwarz, 1978), were found to provide unsatisfactory results. This is mainly due to the use of spectral techniques, since they map the elements to be clustered onto a different eigenspace whose dimensions change with the number of clusters considered, complicating the task of comparing the goodness-of-fit for different number of clusters by such probabilistic methods. Instead, the *silhouette coefficient* (Rousseeuw, 1987), SC, is used. It is a confidence indicator of the membership of an element to the cluster it was assigned to. It is defined, for each element e_i, as $SC_i = (b_i - a_i)/\max(a_i, b_i)$, where a_i is the average distance between element e_i and other elements in its cluster, and b_i is the average distance to the items in the closest cluster. It varies from -1 (lowest membership) to $+1$ (highest membership). Being a dimensionless quantity, the mean and variance throughout all the clustered elements can be used as indicators of the optimality of the clustering, and compared among results for different numbers of clusters to determine the optimum number of them. High values of the mean silhouette indicate a high degree of membership of the elements being clustered to the clusters they were assigned, and low values of its variance indicate that the majority of elements have a similar value of the silhouette coefficient. The combination of both indications reflects a successful clustering.

Once the cluster centers have been obtained, it is also possible to retrieve the closest elements to those cluster centers among the elements being classified. These closest elements to the cluster centers can be considered as representative elements of each cluster.

Chapter 3

Application to a virtual set of structures

The validity and applicability of the last two steps (characterization and classification) of the proposed methodology were tested on a virtual set of nearly 200 surfaces created using computer three-dimensional modeling tools. The extraction step was not included in this test, since the starting point is the set of surfaces itself. Nevertheless, modeled structures of very different sizes were included, to emulate the multiple scales that would result from the extraction step, had it been included. Also, the shapes of the modeled structures are all different. They could be visually classified into three main groups with a common geometry: blobs, tubes, and sheets. The target of this test was to educe those three main groups automatically and without any a priori knowledge of the possible geometries of the structures present in the dataset or of the number of groups among them, that is, simply based on the characterization and classification steps of the methodology previously explained. Among the modeled sheet-like structures, approximately one third were given a certain rolling geometry (spiral-like sheets).

Figure 3.1 shows the visualization space with the results of the test. Each sphere in that space represents a structure of the virtual set (some examples are projected onto the planar sides). The color of each sphere in the visualization space indicates the cluster to which its corresponding structure has been automatically assigned by the clustering algorithm during the classification step of the methodology, and its diameter is scaled using its associated silhouette coefficient, which represents for each structure the degree of membership to the cluster to which it was assigned (refer

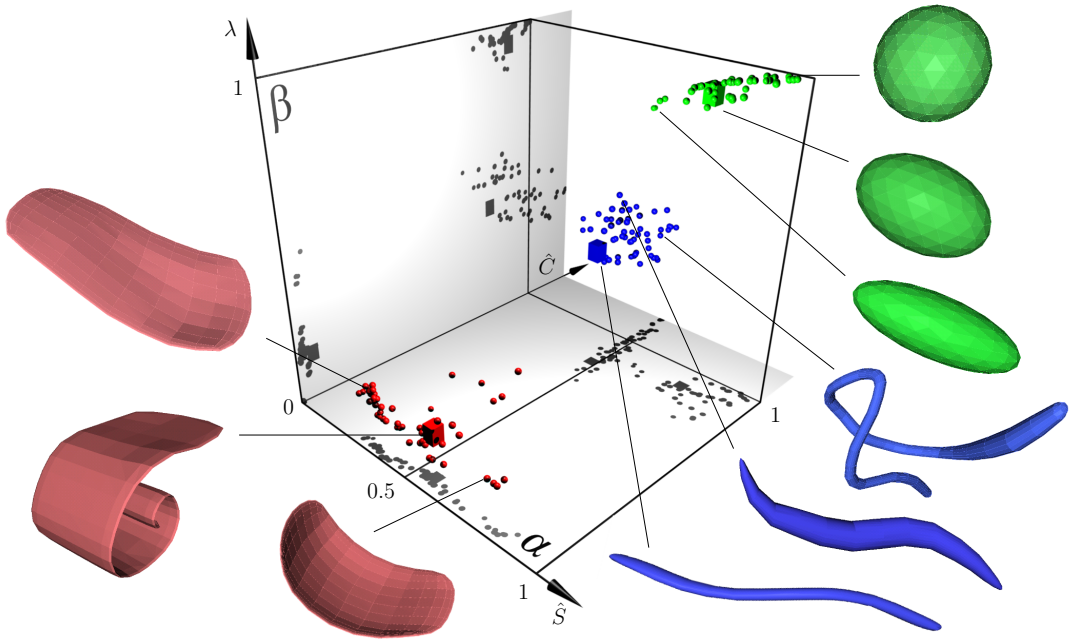


Figure 3.1: Visualization space with clustering results for the virtual set of modeled structures, with representative examples shown at the sides

to §2.3.3), renormalized to have positive values that allow a comparison among structures. For reference, the closest elements to the cluster centers have been highlighted using cubes of slightly bigger size.

Three clusters were automatically educed in the classification step and each structure was 'correctly' assigned by the algorithm to the appropriate group corresponding to the previously constructed geometry. This can be seen in Figure 3.1 from the relative locations of the centers of the glyphs defined by $\hat{S}$, $\hat{C}$, and λ for each structure. We emphasize that neither information on the previously constructed shapes nor the number of groups to be educed formed any part of the clustering algorithm (for example as pre-conditioning). The results of Figure 3.1 are a consequence of the geometric characterization and automatic classification in the feature space of parameters.

We note also that in Figure 3.1 (as was sketched in Figure 2.8), the sheet-like structures can spread over a large region near the plane $\hat{C} = 0$ in the visualization space. This region could be narrowed by means of a transformation of the $(\hat{S}, \hat{C})$-plane to Cartesian coordinates $(\hat{X} = \hat{C}\cos[\pi(\hat{S} - 1/2)], \hat{Y} = \hat{C}\sin[\pi(\hat{S} - 1/2)])$. This would bring sheet-like structures to the axis $(0, 0, \lambda)$ in the new visualization

space. Nevertheless, it is helpful to keep the original visualization and feature spaces, since that allows a possible distinction of the different shapes of the structures that fall into the broadly defined sheet-like geometry. For example, in the test case of the virtual set of modeled structures presented here, the second optimum automatic clustering result was such that four groups were educed: the blob-like and tube-like clusters remained the same as in the optimum case of three clusters described above, but the sheet-like cluster was split into two, with one of these clusters containing a large proportion of structures with a rolling geometry (spiral-like sheets). This also suggests that further post-processing (ideally also automatic) of the educed clusters can be helpful in refining the results.

Chapter 4

Geometry of structures of a passive scalar fluctuation field in stationary isotropic homogeneous turbulence

4.1 DNS database

We use a numerical database obtained from a DNS with 512^3 grid points; the incompressible Navier–Stokes equations for the velocity field and the advection–diffusion equation for the passive scalar fluctuation were solved by means of a Fourier-Galerkin pseudo-spectral method. The domain is a cube of side 2π with periodic boundary conditions. The velocity field was forced at large scales, becoming statistically stationary in time. A mean scalar gradient was imposed so that the scalar fluctuation field became also statistically stationary in time. Despite the mean scalar gradient applied, the scalar fluctuation is statistically homogeneous. The Reynolds number based on the integral length scale is 1901, whereas the Taylor Reynolds number is $Re_\lambda = 265$. The Schmidt number of the simulation is $Sc = 0.7$. The product of the largest dynamically significant wavenumber, $k_{\max}$, and the average Kolmogorov length scale, $\bar{\eta}$, is $k_{\max}\eta = 1.05$. More specific details of the database can be found in O'Gorman & Pullin (2004).

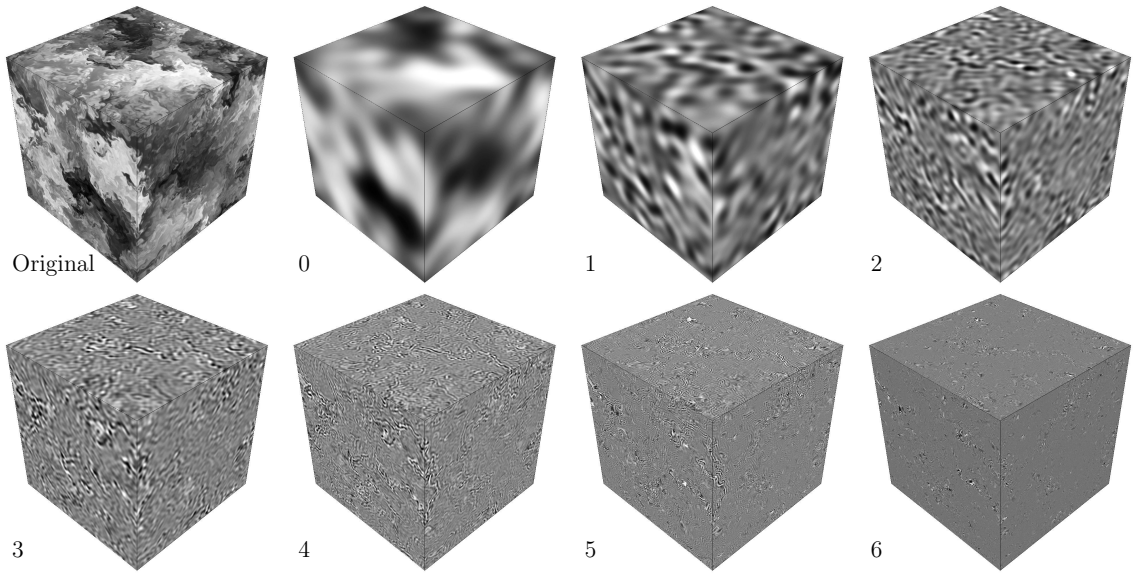


Figure 4.1: Tri-plane cuts of the passive scalar fluctuation field for the original database (top left) and each of the filtered scales resulting from the multi-scale analysis (filtering in curvelet domain) (increasing scale number from left to right and top to bottom)

4.2 Multi-scale diagnostics

We apply our methodology to the three-dimensional scalar field given by the passive scalar fluctuation at an instant in time. For the given resolution of 512^3 grid points, and a coarsest scale $j_0 = 2$, the curvelet transform provides seven scales. They will be named by a scale number, from 0 to 6; increasing values of the scale number correspond to smaller scales. Thus, scale 0 captures the largest scales and 6 the smallest. Figures 4.1 and 4.2 show the result of the multi-scale analysis based on the curvelet transform. Plane cuts of the original database and each of the filtered scales (filtered in the curvelet domain and then inverse transformed to the physical domain) are shown. Three-dimensional views with plane cuts in the three directions of the volume data are presented in Figure 4.1, and more detailed plane cuts normal to the x_3-direction at half the length of the cube are shown in Figure 4.2. Volume-data pdfs obtained for the scalar field associated with the original database and for each one of the filtered scales, Figure 4.3(a), give insight into the distribution of the scalar values at the different scales and their contribution to the total field (original database). For this scalar field, the pdfs tend to become narrower for increasing values of the scale number, that is, for smaller scales. Scalar fluctuation spectra are also computed for the original volume data

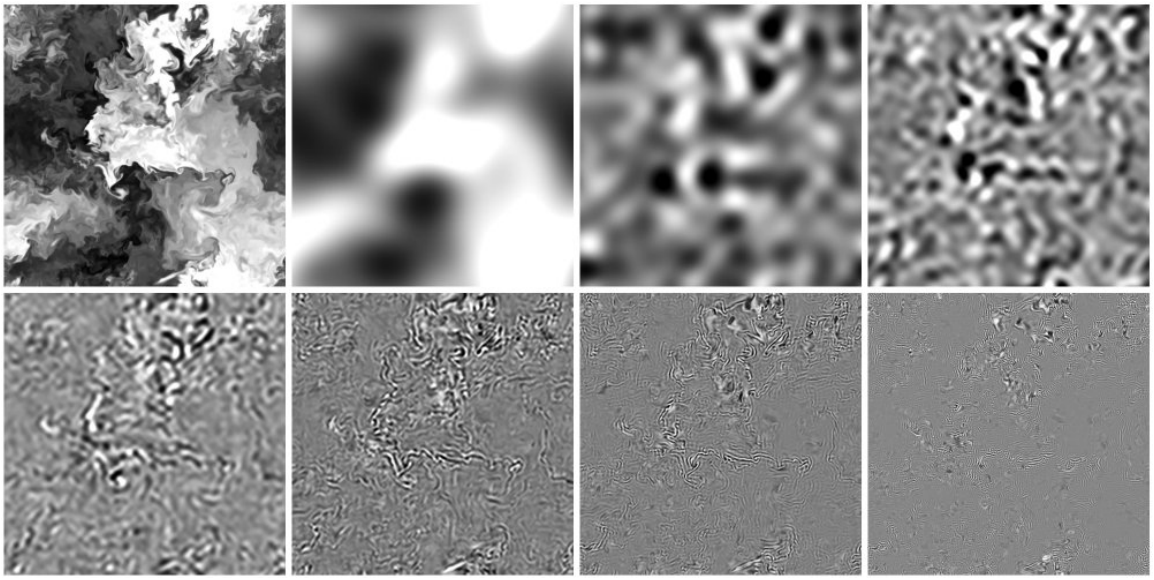


Figure 4.2: Plane cuts normal to the x_3-axis at its midpoint of the passive scalar fluctuation field for the original database (top left) and each one of the filtered scales resulting from the multi-scale analysis (filtering in curvelet domain) (increasing scale number from left to right and top to bottom)

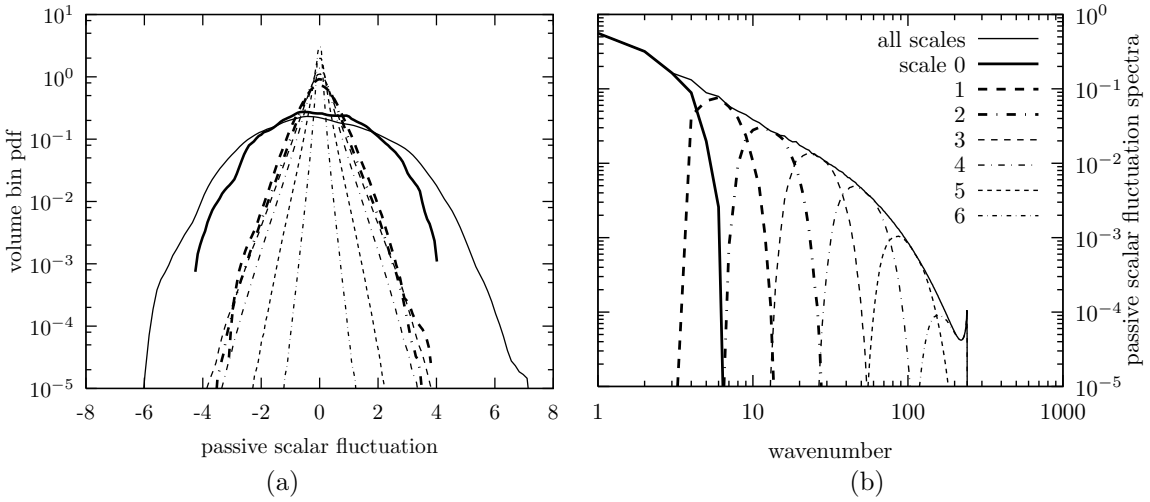


Figure 4.3: Volume-data pdfs of the passive scalar fluctuation field (a) and corresponding spectra (b), associated with the original database (containing all scales) and each of the filtered scales

field and each filtered scale and are shown in Figure 4.3(b). The effect of the curvelet filtering in the Fourier domain can be observed. This differs from a top-hat window filtering in that domain, in order to preserve the localization in the physical domain. It can be noticed that scales 1, 2, and 3 correspond mainly to the inertial range of scales, whereas scales 4 and 5 are mainly dissipation scales. From this observation and from Figure 4.3(a) we note that those pdfs associated with the scales corresponding to the inertial range (1, 2 and 3) are very similar, almost collapsing in that

plot.

Additionally, an equivalent multi-scale decomposition is done for the velocity field. That allows us to define characteristic squared integral velocities, $\overline{u_i^2}$, and integral length scales, L_i, L_i', for each filtered scale, i, in the same terms in which they are defined for the original velocity field. For the case of isotropic turbulence, they can be expressed as

$$\overline{u_i^2} = \frac{2}{3}\int_0^\infty E_i(k)\mathrm{d}k, \tag{4.1}$$

$$L_i = \frac{\pi}{2\overline{u^2}}\int_0^\infty \frac{E_i(k)}{k}\mathrm{d}k, \tag{4.2}$$

$$L_i' = \frac{\pi}{2\overline{u_i^2}}\int_0^\infty \frac{E_i(k)}{k}\mathrm{d}k, \tag{4.3}$$

where $E_i(k)$ is the energy spectrum associated with scale i, and the absence of subindex refers to the original velocity field. It follows from equation 2.5 that the sums of the energy spectra and the characteristic squared integral velocities of all the filtered scales is equal to those of the original velocity field, $E(k)$ and $\overline{u^2}$, respectively:

$$E(k) = \sum_i E_i(k), \tag{4.4}$$

$$\overline{u^2} = \sum_i \overline{u_i^2}. \tag{4.5}$$

Table 4.1 shows, for the original velocity field and for each filtered scale, the characteristic squared integral velocity and the characteristic integral length scale, and how they compare to the total characteristic squared integral velocity, $\overline{u^2}$, and the average Kolmogorov length scale, $\bar{\eta}$.

4.3 Geometry of passive scalar iso-surfaces

After the multi-scale analysis, iso-surfaces are obtained for each of the filtered scales. The contour values are, for each filtered scale, the mean value of the scalar field plus two times the standard deviation of that field (mean and standard deviation values can be obtained from the first- and

scale	$\overline{u_i^2}/\overline{u^2}$	$L_i/\bar{\eta}$	$L_i'/\bar{\eta}$
original	1.000	249.6	249.6
0	0.591	226.9	383.8
1	0.155	14.68	96.1
2	0.113	5.235	46.2
3	0.085	1.927	22.8
4	0.044	0.519	11.9
5	0.011	0.070	6.3
6	0.001	0.004	3.3

Table 4.1: Breakdown of characteristic integral velocities and length scales for the filtered scales

second-order moments of the volume pdfs presented before) (see Figure 4.4). Those iso-surfaces corresponding to the same relative contour value at each scale will be characterized and classified and their results compared among the different scales. We also remark that an additional step in the extraction is applied to periodically reconnect those structures intersecting boundaries with their continuation on the opposite boundaries; this reconnection is performed for each individual filtered scale. Both the largest scale (0) and the smallest scale (6) are not considered in the process: the largest scale is of less relevance in this analysis since its structure is expected to depend on the boundary conditions and external forces applied. The smallest scale is excluded to avoid the extraction of spurious structures and/or an erroneous geometrical characterization that could result from the lack of grid resolution, or aliasing effects at that scale. Thus, the scale numbers under analysis are 1–5. In the same spirit, a minimum number of points (300) was considered for a structure to be analyzed, so that it is smooth enough for a reliable calculation of its differential-geometry properties, the basis of the characterization step.

Then, each structure is geometrically characterized as described in §2.2 and, based on the parameters extracted from its signature, it can be represented in the visualization space referred in §2.3. Figures 4.5 and 4.6 show the distribution of glyphs representing each structure, for the different scale numbers 1–5 considered. In this case, glyphs are spheres whose centers correspond to the $\hat{S}, \hat{C}, \lambda$ parameters and whose radii are scaled according to the surface area of the structure, and their color is assigned based on the scale number to which the structure belongs. First, structures of all scales are shown and then the progression for individual scales is presented. As can be seen, the structures

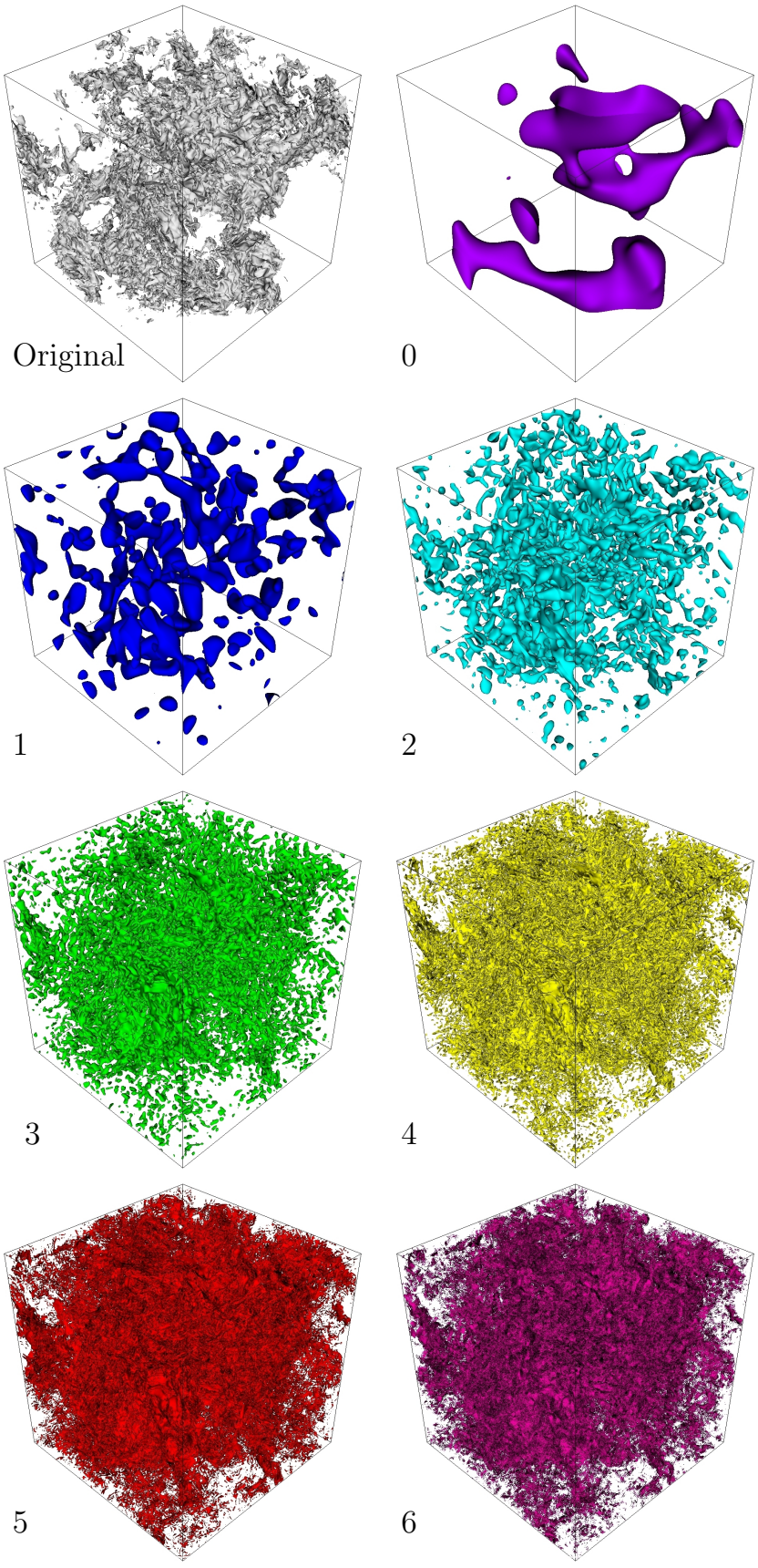


Figure 4.4: Iso-contours of the passive scalar fluctuation field for the original database (top left) and each one of the filtered scales resulting from the multi-scale analysis (filtering in the curvelet domain) (increasing scale number from left to right and top to bottom). A contour value equal to the mean plus two times the standard deviation of each resulting scalar field was used

go from predominantly blob-like and tube-like at scales 1–3 toward more sheet-like structures at the smaller scales 4 and 5, The stretching of the structures increases with the scale number; that is, the parameter λ decreases for smaller and smaller scales.

Some representative structures, named A–L, have been selected (see top of Figure 4.5); their corresponding signatures are shown in Figure 4.7. The cascade in the passive scalar fluctuation spectrum is thus translated into a cascade of the representation of structures in the visualization space, that starts near the point $(1, 1, 1)$ (sphere) and evolves toward highly stretched sheet-like structures. whose geometry tends to be complex (see, for example, their corresponding signatures in the last few points in Figure 4.7).

The clustering algorithm is then applied to the structures. The set of parameters $\{\hat{S}, \hat{C}, \lambda, d_l^S, d_u^S, d_l^C, d_u^C\}$ is used to form the feature space where each structure is represented by a point. Three groups of structures are obtained, and the result can be seen in Figure 4.8. That figure shows a visualization space with the same three spatial coordinates as in earlier plots (S, C, λ), and the structures represented by glyphs consisting of spheres (colored by the cluster ID and with radius scaled by the silhouette coefficient, defined in §2.3.3, that indicates the level of membership to the cluster to which it has been assigned) and horizontal bars with origin at the center of the sphere and lengths proportional to the other four parameters used for clustering (distances d_u and d_l in $\pm S$ and $\pm C$ directions). The thickness of these bars is also scaled by the value of the silhouette coefficient. Although the clustering algorithm captures the main trends, the structure geometries appear continuously distributed across the main groups, rather than separated into well-differentiated groups. This translates into the distribution of points and glyphs representing structures in the feature and visualization spaces: in particular, glyphs associated with the educed structures are organized as a cloud in the visualization space, transitioning from one region to other regions. For example, a comparison of the clustering results for the passive scalar field (Figures 4.8) and the test case of modeled structures (Figure 3.1) previously presented in §3 clearly shows the difference between the continuously distributed geometries of the structures educed for the passive scalar field, and the well distinct groups of geometries found in the test case. As a result, for the case of the passive scalar,

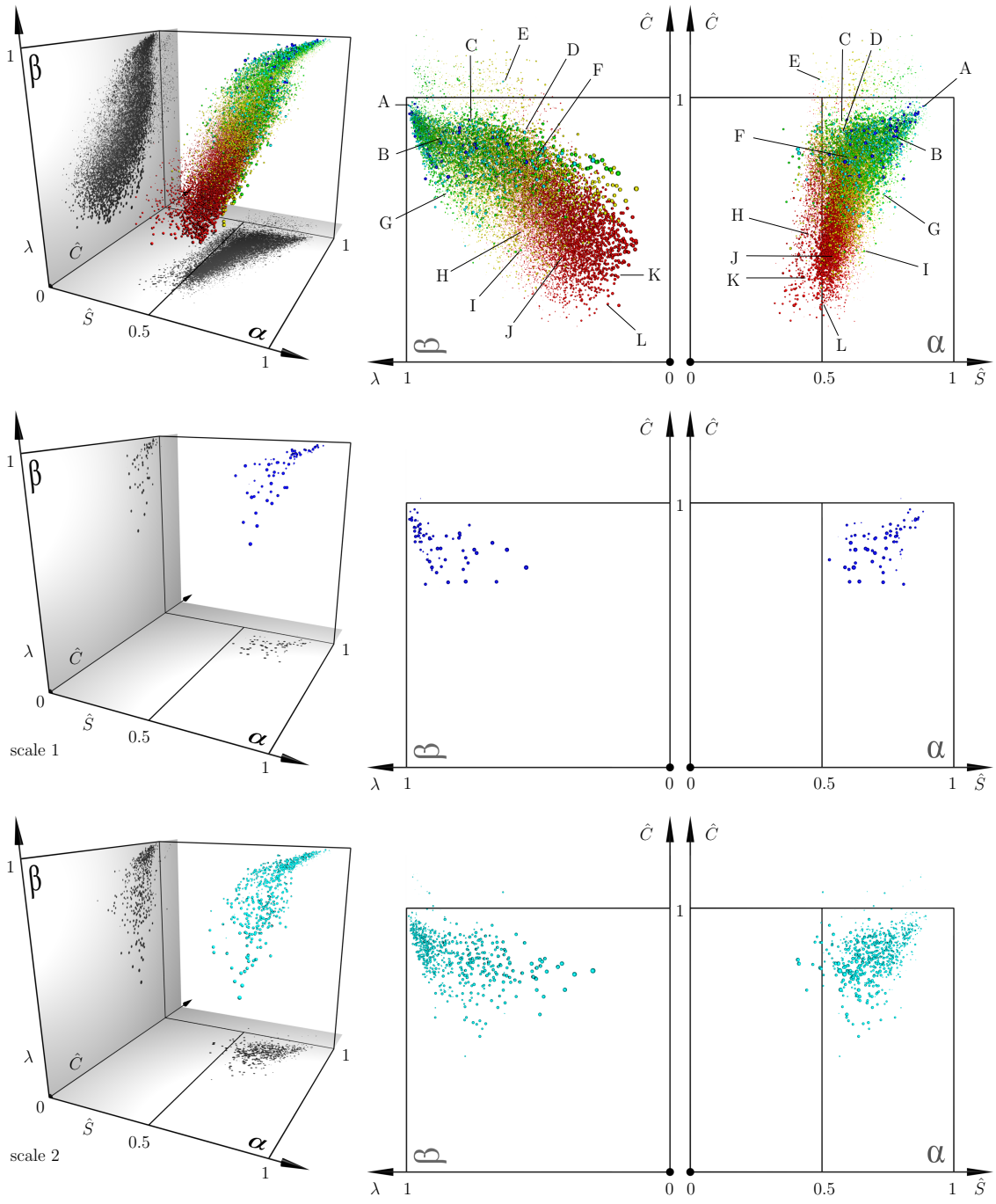


Figure 4.5: Visualization space with spheres representing the structures educed from the passive scalar fluctuation field at all scales (top), and only at scale 1 (middle) and 2 (bottom). Radii of spheres represents (in normalized logarithmic scale) the surface area of each structure. Color of the spheres represents the scale to which the structure belongs: dark blue (1), light blue (2), green (3), yellow (4), red (5). (Continued on Figure 4.6)

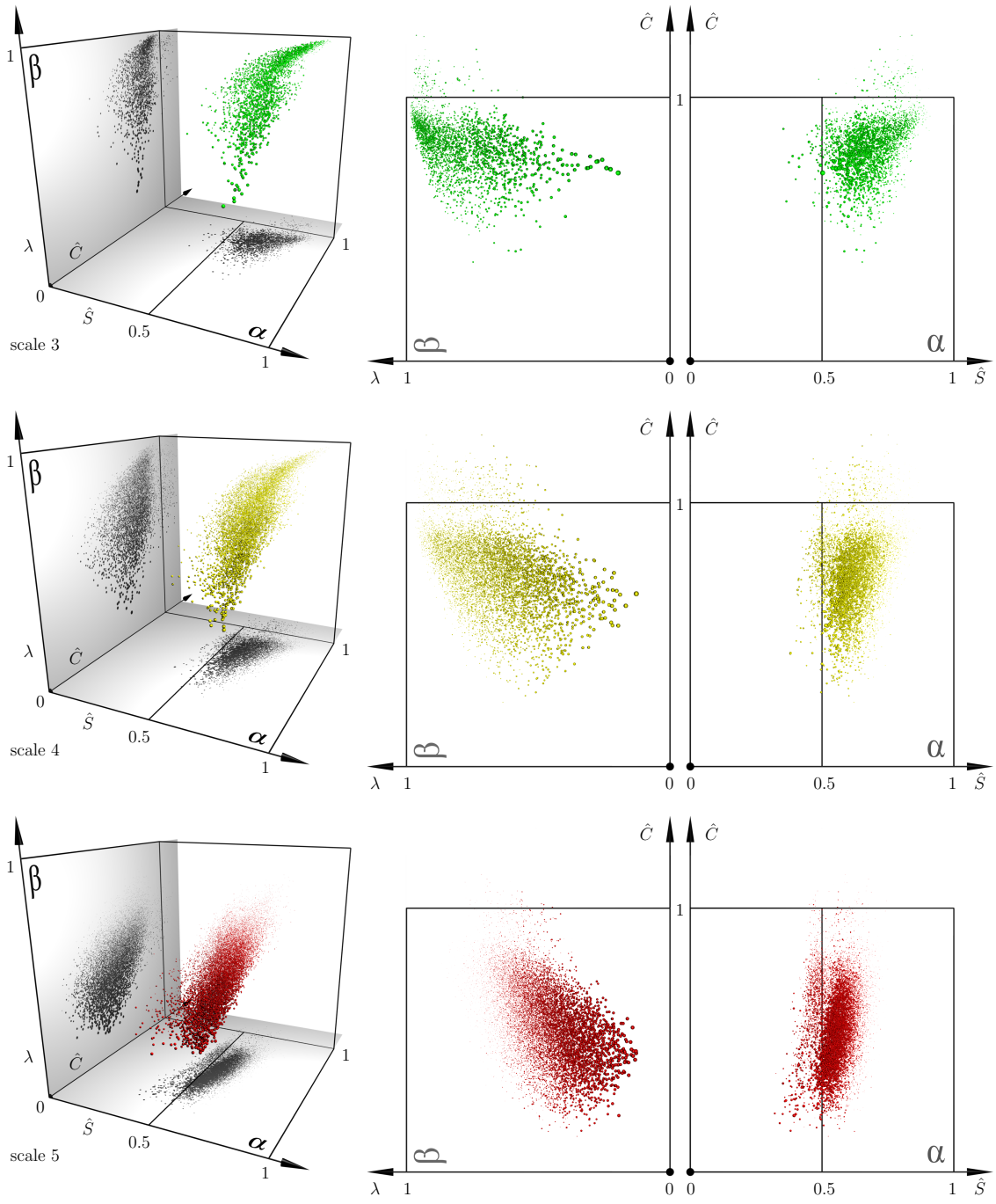


Figure 4.6: Visualization space with spheres representing the structures educed from the passive scalar fluctuation field at scale 3 (top), 4 (middle), and 5 (bottom). Radii of spheres represents (in normalized logarithmic scale) the surface area of each structure. Color of the spheres represents the scale to which the structure belongs: green (3), yellow (4), red (5). (Continued from Figure 4.5)

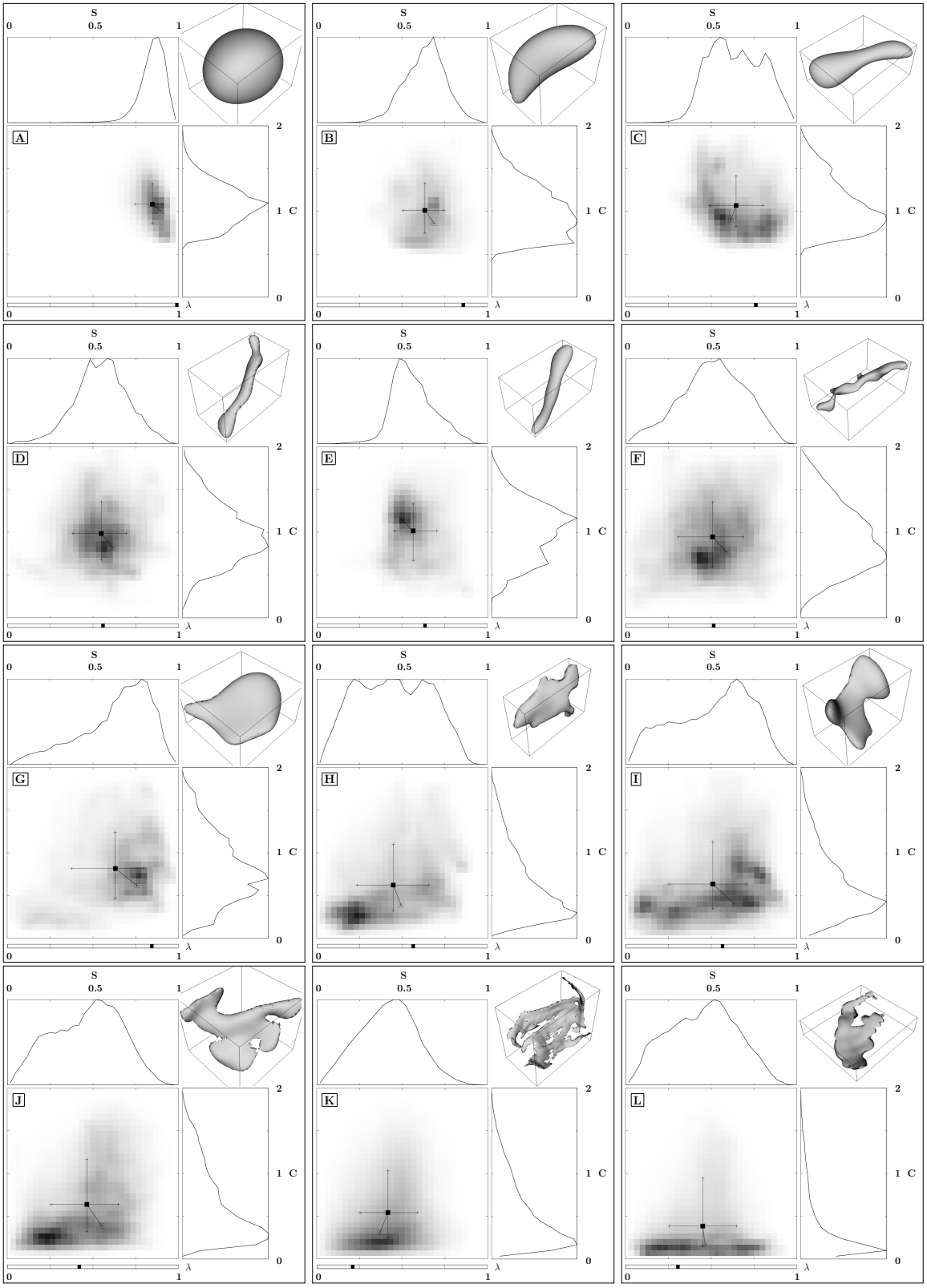


Figure 4.7: Signatures of representative structures (refer to Figure 4.5 (top) for the location of the corresponding points (A–L) in the visualization space)

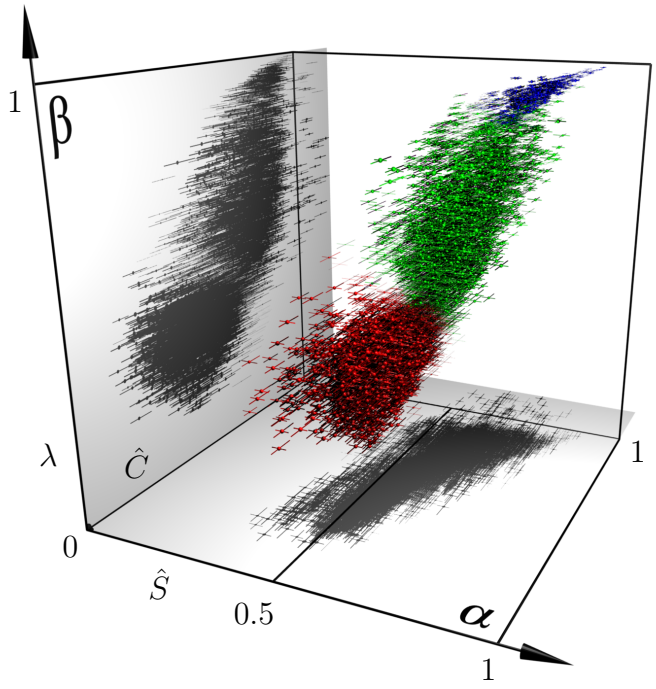


Figure 4.8: Result of the clustering algorithm in a visualization space showing spheres representing structures with radii scaled by the value of their silhouette coefficient (renormalized to have only positive values) and colored by the cluster to which they belong. An optimum number of three clusters was automatically found. The seven clustering parameters used to define the feature space are shown for each structure by the center of its representing sphere (of coordinates $\hat{S}, \hat{C}, \lambda$) and the four bars scaled by the value of d_u, d_l of S and C

some of the structures classified as belonging to one group but lying on the overlapping regions in the feature space will not necessarily be significantly different (geometrically) from other structures that belong to other groups but with a similar location in the feature space. The degree of membership to the educed clusters (measured by the silhouette coefficient) of those structures in the overlapping regions between clusters will therefore be lower than that of structures near the cluster centers. Note how the glyphs in Figure 4.8, scaled by the renormalized silhouette coefficient of the associated structures, are smaller in the overlapping regions (compare, for example, Figure 4.5 (top), where the density of points is much more continuous throughout the whole distribution, since the scaling factor in that case was the area of the structure, not its silhouette coefficient).

4.4 Discussion and physical interpretation

We discuss first the smallest scales. Figure 4.6 (bottom) shows that highly stretched sheet-like structures are predominant at those scales. From the size of the spheres representing the structures (related to the the surface area of the structures in a normalized logarithmic scale), we conclude that the larger of these structures appear more stretched (lower λ) and tend to be more sheet-like (lower $\hat{C}$). This trend is in agreement with the results of Schumacher & Sreenivasan (2005), who, using a conventional box-counting method, found that passive scalar iso-level sets at the smallest scales become smooth sheets. They did not find fractal structure in the passive scalar field within their range of Sc and Re_λ, which differ from present values. Schumacher & Sreenivasan also used the area-to-volume ratio of the iso-levels, which is essentially the parameter μ used in our methodology to form the dimensionless curvedness $\hat{C}$: they use this in a global sense rather than applied to individual structures, as is done here. In the study of intense strain structures in homogeneous isotropic turbulence by Moisy & Jiménez (2004), the dominance of sheet-like structures in the smallest scales, suggested by a similar box-counting method, was confirmed by the geometrical study of aspect ratios of individual structures.

Multi-scale decompositions of the vorticity field in turbulent flows have been previously applied using orthogonal wavelets in two and three dimensions (see Farge et al., 1999, 2001, 2003), where thresholding of the wavelet coefficients based on denoising theory separates the vorticity into two orthogonal fields, denoted as *coherent* and *incoherent* vorticity. It is found that the coherent field is responsible for most of the energy transfer in the large and inertial scales. In three-dimensional homogeneous isotropic turbulence, the pdf of the coherent vorticity is found to be stretched exponential while the incoherent vorticity is exponential. We find that the pdf of the scalar fluctuation (see Figure 4.3) is Gaussian, with sub-Gaussian tails, in agreement with previous results for scalar fields (Overholt & Pope, 1996; Celani et al., 2001). In our multi-scale decomposition (see Figure 4.3), the variance of the pdfs also decreases for smaller scales (resulting in narrower pdfs), quickly transitioning from Gaussian (with slightly sub-Gaussian tails) to exponential. Presently our multi-scale analysis does not include any assumptions about the 'coherence' of the educed structures through

thresholding of the multi-scale coefficients, since it has a different purpose, mainly as a diagnostics tool.

It has been suggested that exponential tails of the scalar fluctuation pdfs are linked to metrics of anomalous mixing (see the discussion in Warhaft (2000)). There have been attempts to clarify the conditions under which sub-Gaussian/exponential tails appear based on various factors that include relative simulation box size (Overholt & Pope, 1996; Schumacher & Sreenivasan, 2005), Reynolds number (distinguishing between 'soft' and 'hard' turbulence (Jayesh & Warhaft, 1992)), flow initial conditions and forcing (Jaberi et al., 1996), and, for the vorticity, the structure of intense portions of the field (Siggia, 1981; Kerr, 1985; Jiménez et al., 1993). The transition, seen in Figure 4.3, from Gaussian pdf in the larger scales to predominantly exponential pdfs in the smaller scales may be related to the geometry of individual structures present at each scale.

The presence of ramp-cliff structures in the scalar field (plateau–cliff in the scalar fluctuation) (Antonia et al., 1979) has been associated with anisotropy of the passive scalar field in the presence of a mean gradient (Celani et al., 2001; Overholt & Pope, 1996; Warhaft, 2000). These features are seen in the large scales (plateau regions) as well as in the small scales (cliffs of fronts), where steep changes in the values of the passive scalar occur (see Figure 4.2). The highly stretched, sheet-like structures found here at the smaller scales could be related to these fronts. A study of their spatial distribution with respect to proximity to the structures of the larger scales would be needed to confirm this. Further, the predominant orientation of the sheet-like regions of such structures could help to clarify their role. This could be obtained within the framework of the present methodology using the multi-orientation decomposition of the curvelet transform (not applied in this thesis).

For the present Sc and Re we know of no previous reports of blob- and tube-like (with moderate stretching) structures in the intermediate scales of the passive scalar fluctuation field. Theoretical developments in physical models of passive scalar mixing have utilized tube-like structures, stretched by large-scale strain fields, to analyze cascade and dissipation dynamics for a passive scalar (Pullin & Lundgren, 2001). Small-scale scalar mixing is modeled as a two-dimensional blob (a tube in three dimensions) convecting, deforming, and diffusing in the presence of the swirling motion of

a stretched spiral vortex. The blob is drawn out into rolled-up sheets whose azimuthally averaged structure remains tube-like. Within this compound tube-sheet structure, the derived scalar spectrum comprises two parts in the form of Batchelor (1959) k^{-1} and Obukov-Corrsin $k^{-5/3}$ (see Tennekes & Lumley (1974)) components. The $k^{-5/3}$ contribution arises from the non-axisymmetric scalar field which tends to be sheet-like. This dynamical model is not inconsistent with the present findings of tube and sheet structures at the smallest scales. We can hypothesize that blob-like structures, similar to structure A shown in Figure 4.7, are created first. These are then strained and stretched by the action of vortex tubes (D, E, F) to form vortex sheets (J, K, L). Further vortex tubes are then created by rolling-up of the sheets. Additional support for the validity of this picture as a physical mechanism of the cascade would probably require (at least) local correlation in tube and sheet structure locations and perhaps orientations for adjacent scales in the sense of the curvelet transform. Some of these topics will be covered in Chapter 7.

Chapter 5

Geometry of structures of enstrophy and dissipation fields in decaying homogeneous isotropic turbulence

5.1 DNS database

The numerical database used here is that of Horiuti & Fujisawa (2008), henceforth referred to as HF. It corresponds to a DNS of incompressible homogeneous isotropic turbulence decaying in time in a cubic domain of side length 2π, with periodic boundary conditions. We use runs corresponding to 256^3, 512^3, and 1024^3 grid points with the same value of the kinematic viscosity, ν, resulting in a similar Taylor-microscale Reynolds number, $Re_\lambda \approx 77$, and in grid resolution criteria, $k_{\max}\bar{\eta}$, of approximately 1, 2, and 4, respectively. Together with the fact that the initial conditions are the same for the three runs, they can be used to compare the geometry of flow structures at different resolutions. Additional parameters for each run at the instant of maximum enstrophy are shown in Table 5.1, extracted from HF, where more details of the computational method can be obtained. The three velocity fields at this time instant for the three grid resolutions are the database for the present study.

In this chapter, we apply the methodology proposed in Chapter 2 for the study of the geometry of structures in turbulence to the enstrophy and dissipation fields. Local enstrophy is defined as

N	ν	Re_λ	$\langle K \rangle$	$\langle \epsilon \rangle$	L	λ	$\bar{\eta}(\times 10^{-3})$	$k_{\max}\bar{\eta}$
256^3	0.00138	77.20	0.900	0.654	0.469	0.138	8.00	1.02
512^3	0.00138	76.87	0.897	0.654	0.469	0.137	8.00	2.05
1024^3	0.00138	77.43	0.904	0.654	0.468	0.138	8.00	4.09

Table 5.1: Parameters for the computed cases: grid points, N; kinematic viscosity, ν; Taylor-microscale Reynolds number, Re_λ; average kinetic energy, $\langle K \rangle$; average dissipation rate, $\langle \epsilon \rangle$; integral length scale, L; Taylor microscale, λ; average Kolmogorov length scale, $\bar{\eta}$; grid resolution criterion, $k_{\max}\bar{\eta}$ (where $k_{\max}$ is the largest dynamically significant wavenumber). From HF

$\omega_i\omega_i$, while local dissipation is defined as $\epsilon \equiv 2\,\nu S_{ij}S_{ij}$. S_{ij} is the strain-rate tensor and ω_i is the vorticity field, which is related to the rotation-rate tensor, Ω_{ij}, by $\Omega_{ij} = -\epsilon_{ijk}\,\omega_k/2$, ϵ_{ijk} being the Levi-Civita symbol. Thus: $\omega_i\omega_i = 2\,\Omega_{ij}\Omega_{ij}$. Strain- and rotation-rate tensors are obtained from the velocity gradient tensor, $\partial u_i/\partial x_j$, as:

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad \Omega_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right). \tag{5.1}$$

We will consider the dissipation rescaled by $(2\,\nu)^{-1}$, that is, $S_{ij}S_{ij}$.

5.2 Multi-scale decomposition

We plot in Figure 5.1 the volume pdfs (left) and the spectra (right), in Fourier space, of the two fields, $\omega_i\omega_i$ and $S_{ij}S_{ij}$, for the three grid resolutions (256^3, 512^3, 1024^3). It can be observed that the pdf of $\omega_i\omega_i$ has longer tails than that of $S_{ij}S_{ij}$. This indicates that large amplitude events in the enstrophy field are more prevalent than in the dissipation field (see Chen et al., 1997), and thus, that $\omega_i\omega_i$ is more intermittent than $S_{ij}S_{ij}$, which is in agreement with results from experiments (Zeff et al., 2003) and numerical simulations (Siggia, 1981; Kerr, 1985; Chen et al., 1997), using the flatness of each field to measure its intermittency. The discrepancy in the spectra of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ increases with the wavenumber, k, that is, for smaller scales. This is also in agreement with the results from numerical simulations at higher Reynolds numbers of Ishihara et al. (2003), who found that the maximum difference between spectra of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ peaks at $k\bar{\eta} \approx 0.4$.

Figure 5.2 shows the effect in physical space of the multi-scale decomposition for $\omega_i\omega_i$ with 512^3

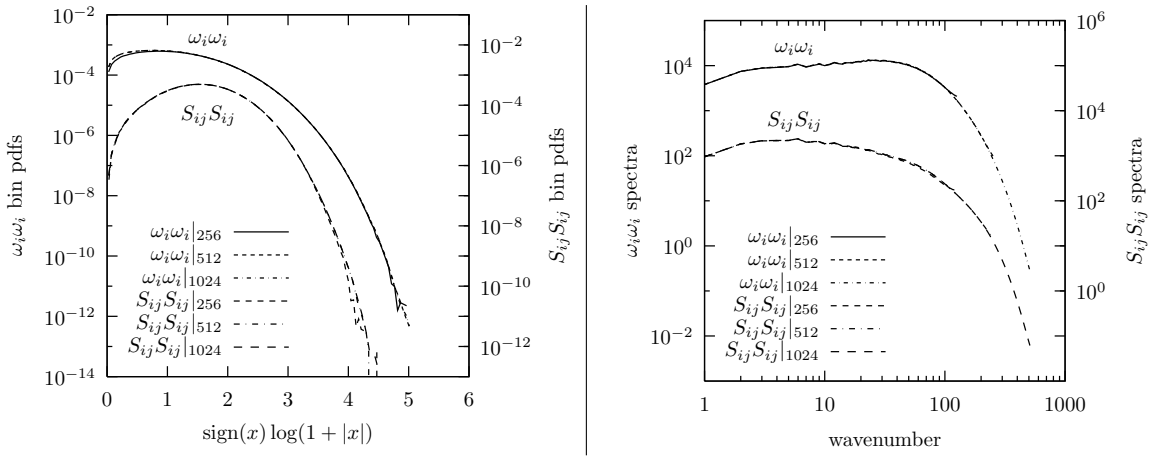


Figure 5.1: Volume pdfs in physical domain (left) and spectra in Fourier domain (right) of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ fields for the three grid resolutions (256^3, 512^3, and 1024^3). Note that the volume pdfs use a transformation of the form $\mathrm{sign}(x)\log(1+|x|)$ in the abscissa coordinate, and that curves for $\omega_i\omega_i$ and $S_{ij}S_{ij}$ fields use two different vertical axes (both in the pdfs and the spectra), shifted one decade for a clear view (non-intersecting curves)

grid points, as an example, through plane cuts in the three principal directions of the cubic domain, compared to the original field (top left).

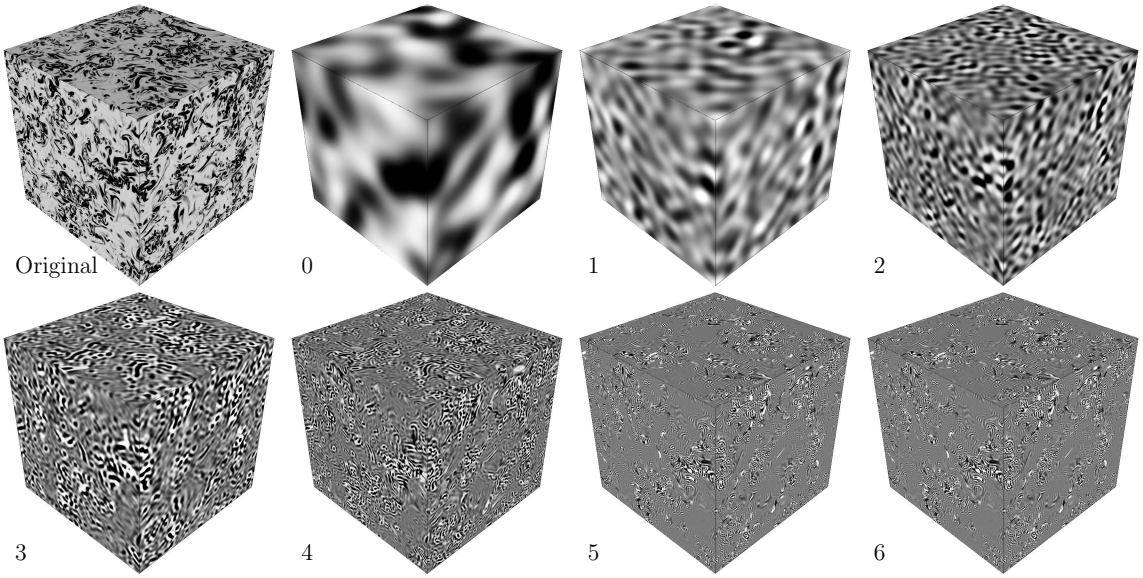


Figure 5.2: Tri-plane cuts of $\omega_i\omega_i$ and its multi-scale component fields for the 512^3 case

Volume pdfs (physical domain) and spectra (Fourier domain) of the original and component fields after the multi-scale decomposition are shown in Figure 5.3, for $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) fields in the 1024^3 case. Scales are named by scale numbers from 0 to 7. Increasing scale numbers, when

referred to the component fields in which the original field is decomposed, indicate smaller scales. Thus, 0 corresponds to the largest scale, and 1, 2,... correspond to smaller and smaller scales. For the 256^3 and 512^3 cases similar plots (not shown) can be obtained with 5 and 6 as the maximum scale numbers respectively. Note how, for both $\omega_i\omega_i$ and $S_{ij}S_{ij}$, the range of the pdfs increases for

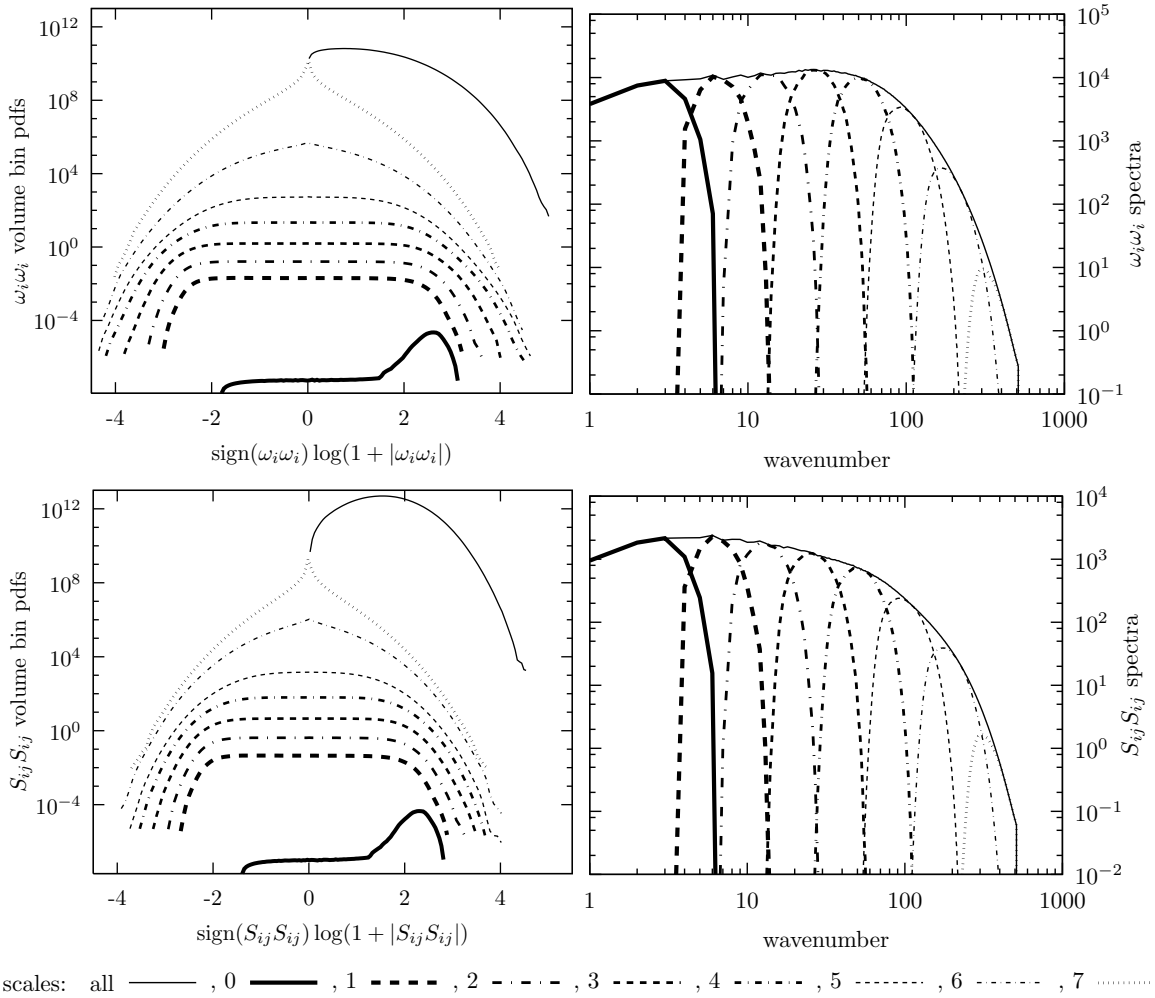


Figure 5.3: Effect of the multi-scale decomposition in the 1024^3 case for $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) fields on the volume pdfs in physical domain (left) and on the spectra in Fourier domain (right). Note that the volume pdfs have been shifted vertically to accommodate all scales and the original fields in a clearer view. Also, instead of using a log-scale in the abscissa of the pdf plots, since there are negative values for all filtered scales, a transformation of the form $\mathrm{sign}(x)\log(1+|x|)$ is used for each field x

increasing scale number (i.e., smaller scales), indicating that fluctuations of both fields are higher in the small scales, and therefore, that intermittency also increases for those smaller scales. Previous multi-scale studies of turbulence have shown this property (Kennedy & Corrsin, 1961; Meneveau,

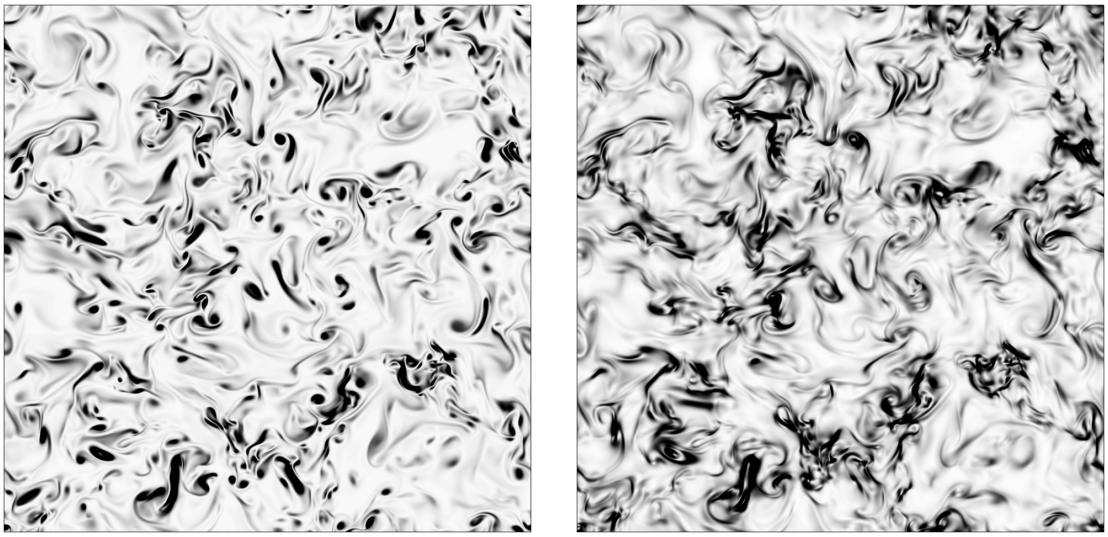


Figure 5.4: Plane cuts of $\omega_i\omega_i$ (left) and $S_{ij}S_{ij}$ (right) normal to one of the principal directions of the cubic domain at half its side length for the 1024^3 case

1991; Brasseur & Wang, 1992; Okamoto et al., 2007).

Figure 5.4 includes plane cuts of $\omega_i\omega_i$ (left) and $S_{ij}S_{ij}$ (right) fields, for the 1024^3, at half the length of the physical domain in one of the principal directions of the cube. Zoomed parts of those plane cuts are shown in Figure 5.5 for the three grid resolutions (256^3, 512^3, and 1024^3, from left to right) for both $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) fields. It can be observed that, particularly in the 256^3 case, the smallest scales are quite different from the higher-resolution cases. Figure 5.6 shows zoomed parts of the plane cuts corresponding to the component field at scale number 5 for the three grid resolutions, which is the highest scale number attainable in the multi-scale decomposition of the 256^3 case and therefore contains the structures at the smallest scales captured in this flow at that grid resolution. It is clear from the two-dimensional fields that structures educed with the lowest grid resolution, 256^3, can be significantly different from the ones at higher grid resolutions, 512^3 and 1024^3. Intuitively, a geometrical characterization of those structures would be affected by that fact, and its effect would be noticed in the application of the methodology proposed above.

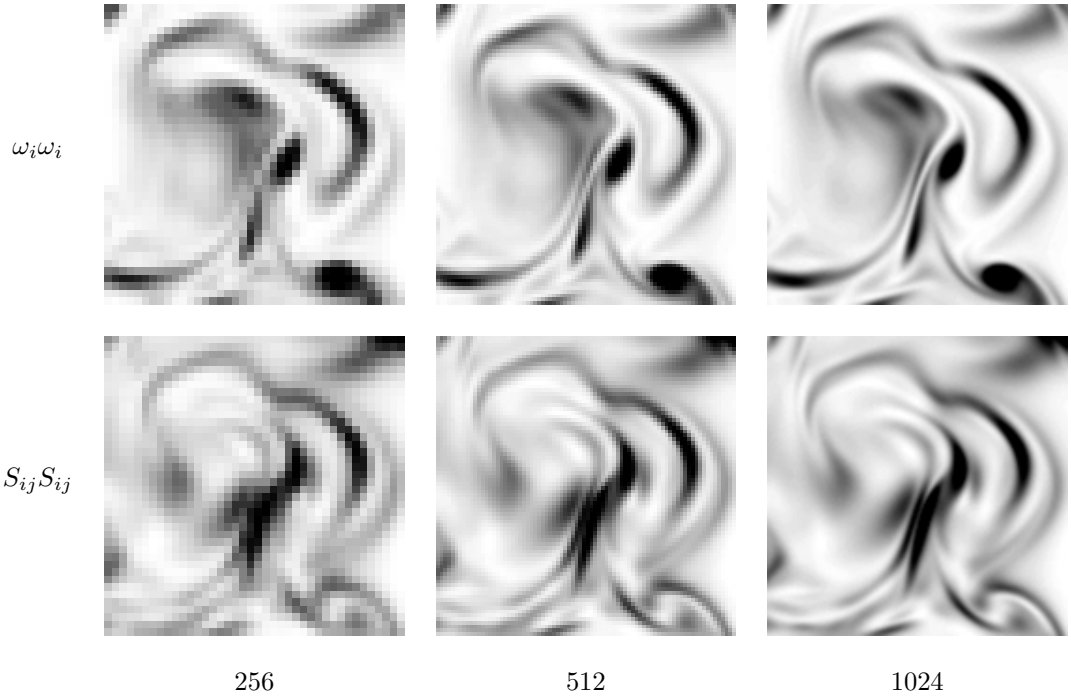


Figure 5.5: Zoomed parts of plane cuts of $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) in one of the principal directions of the cubic domain at half its side length for the three grid resolutions 256^3 (left), 512^3 (center), and 1024^3 (right). Greyscale has been renormalized to the zoomed region for better clarity

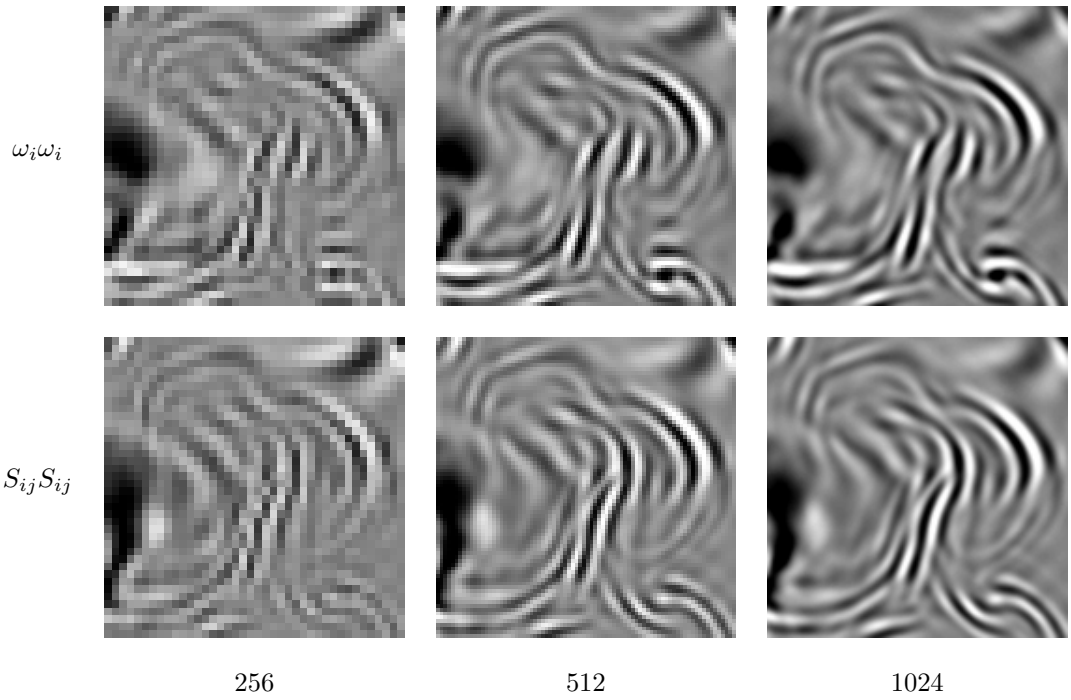


Figure 5.6: Zoomed parts of plane cuts of component field at scale number 5 for $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) in one of the principal directions of the cubic domain at half its side length for the three grid resolutions 256^3 (left), 512^3 (center), and 1024^3 (right). Greyscale has been renormalized to the zoomed region for better clarity

5.3 Characterization and classification of individual structures

In Figure 5.7 we present the three-dimensional visualization spaces (formed by $\hat{S}, \hat{C}, \lambda$ axes) with the glyphs (simple spheres, in this case) representing each structure after its geometrical characterization. The top row corresponds to $\omega_i\omega_i$ and the bottom row to $S_{ij}S_{ij}$, for increasing grid resolution (256^3, 512^3, and 1024^3) from left to right. The spheres are scaled by the lognormalized area of the corresponding structure. The color of each sphere represents the scale number to which it belongs. As in §4.3, the largest scale is not included in the analysis, for being strongly dependent on the boundary conditions and forcing applied. Neither is the smallest scale (for each grid resolution), to avoid interference between grid resolution effects and the iso-contouring process. The fact that we have three different grid-resolutions for the equivalent field allows us to verify whether that interference occurs. This is discussed at the end of §5.4. Thus, scales 1–4, 1–5 and 1–6 are represented for the 256^3, 512^3, and 1024^3 grid resolutions, respectively. Top views ($(\hat{S}, \hat{C})$-plane) of these visualization spaces are shown in Figure 5.8, where the differences between $\omega_i\omega_i$ and $S_{ij}S_{ij}$ fields can be better realized.

In Figure 5.9 we show the breakdown by increasing scale number (top to bottom) of the three-dimensional visualization spaces for the $\omega_i\omega_i$ field for the three grid resolutions (256^3, 512^3, and 1024^3), increasing from left to right. Top views ($(\hat{S}, \hat{C})$-plane) of each visualization space can be seen in Figure 5.10. Figures 5.11 and 5.12 are the equivalent ones for $S_{ij}S_{ij}$.

We discuss first the case with the highest resolution (1024^3). Structures of both $\omega_i\omega_i$ and $S_{ij}S_{ij}$ fields show a continuous transition of their corresponding glyphs in the visualization space, with varying scale (see right plots of Figures 5.7 and 5.8). Structures at the largest scale of both fields (top-right corner of Figures 5.9 and 5.11) are mainly blob-like; some get closer to the tube-like region with small stretching (high λ). At the smallest scale (bottom right corner of Figures 5.9 and 5.11), dominant structures of both fields are sheet-like (low values of $\hat{C}$ and λ). The intermediate scales present a different behavior for each field: $\omega_i\omega_i$ shows a high concentration of structures near the

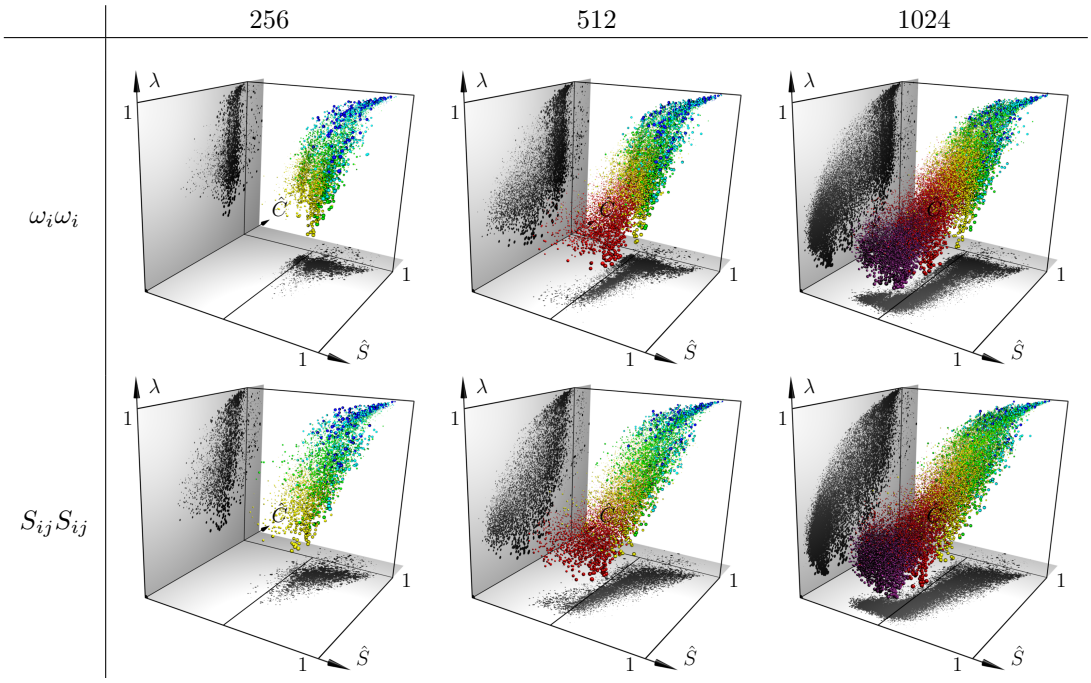


Figure 5.7: 3D views of the visualization spaces, with glyphs (spheres) representing educed structures, colored by scale number (all merged in each visualization space) and scaled by the lognormalized area of the corresponding structure, for $\omega_i\omega_i$ (top row) and $S_{ij}S_{ij}$ (bottom row) at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

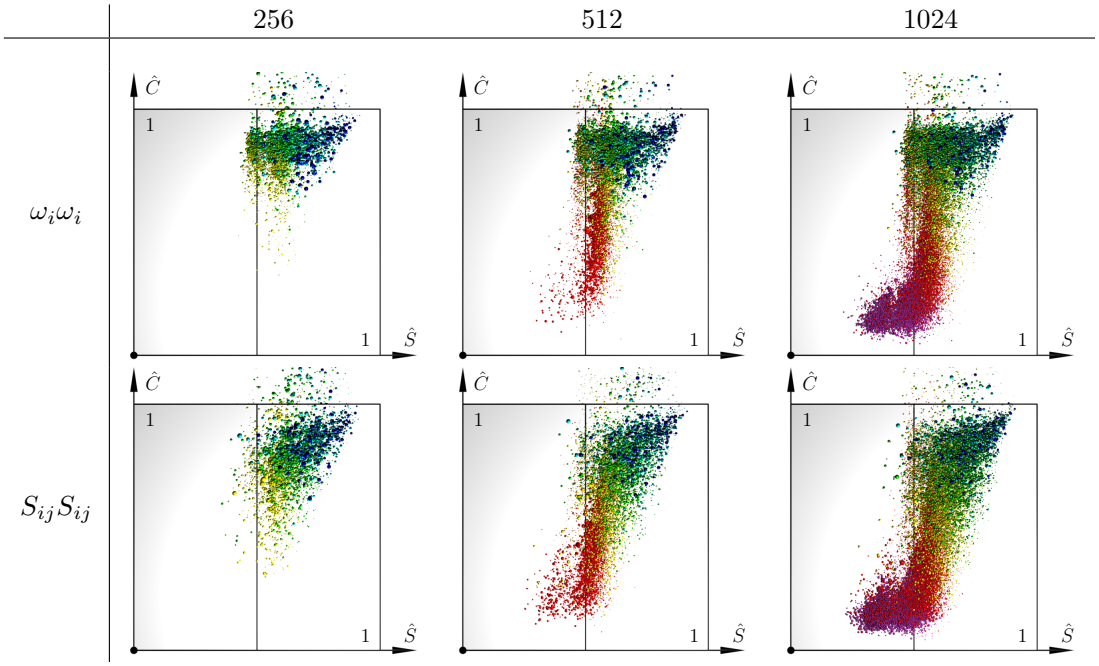


Figure 5.8: Top views ($(\hat{S},\hat{C})$-plane) of the visualization spaces, with glyphs (spheres) representing educed structures, colored by scale number and scaled by the lognormalized area of the corresponding structure, for $\omega_i\omega_i$ (top row) and $S_{ij}S_{ij}$ (bottom row) at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

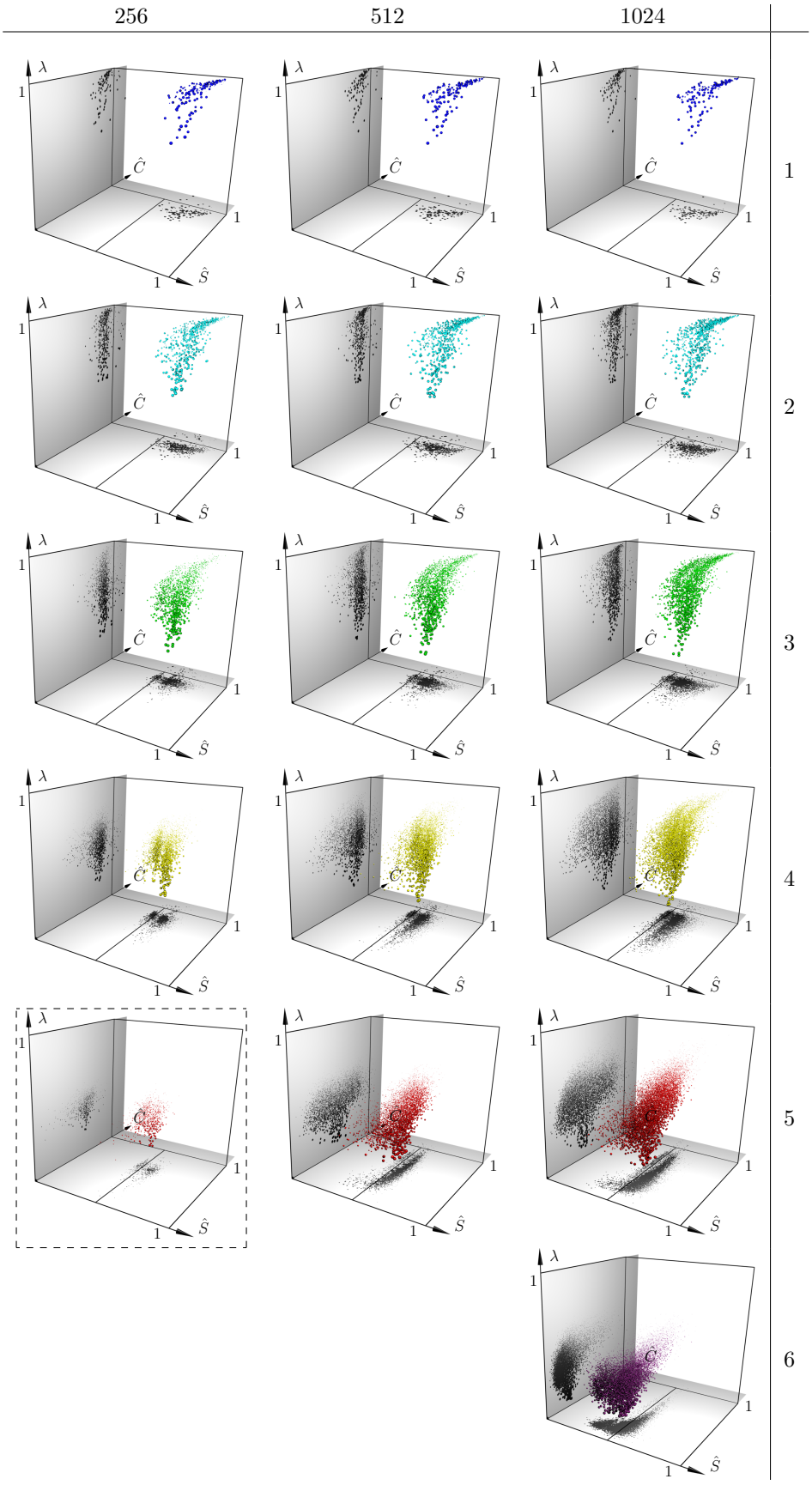


Figure 5.9: Breakdown, by scale number (increasing top to bottom), of 3D views of the visualization spaces for $\omega_i\omega_i$ at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

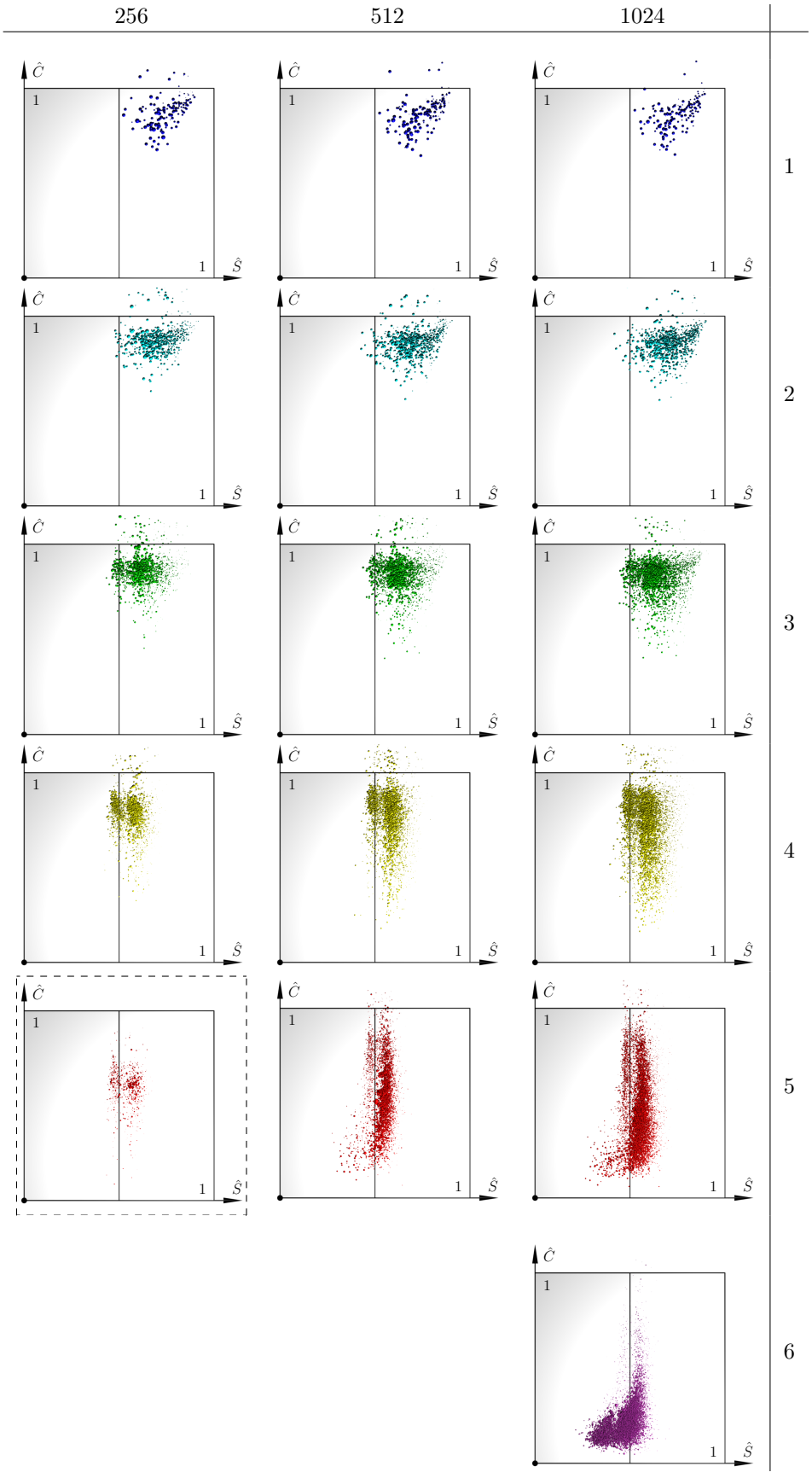


Figure 5.10: Breakdown, by scale number (increasing top to bottom), of top views ($(\hat{S}, \hat{C})$-plane) of the visualization spaces for $\omega_i\omega_i$ at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

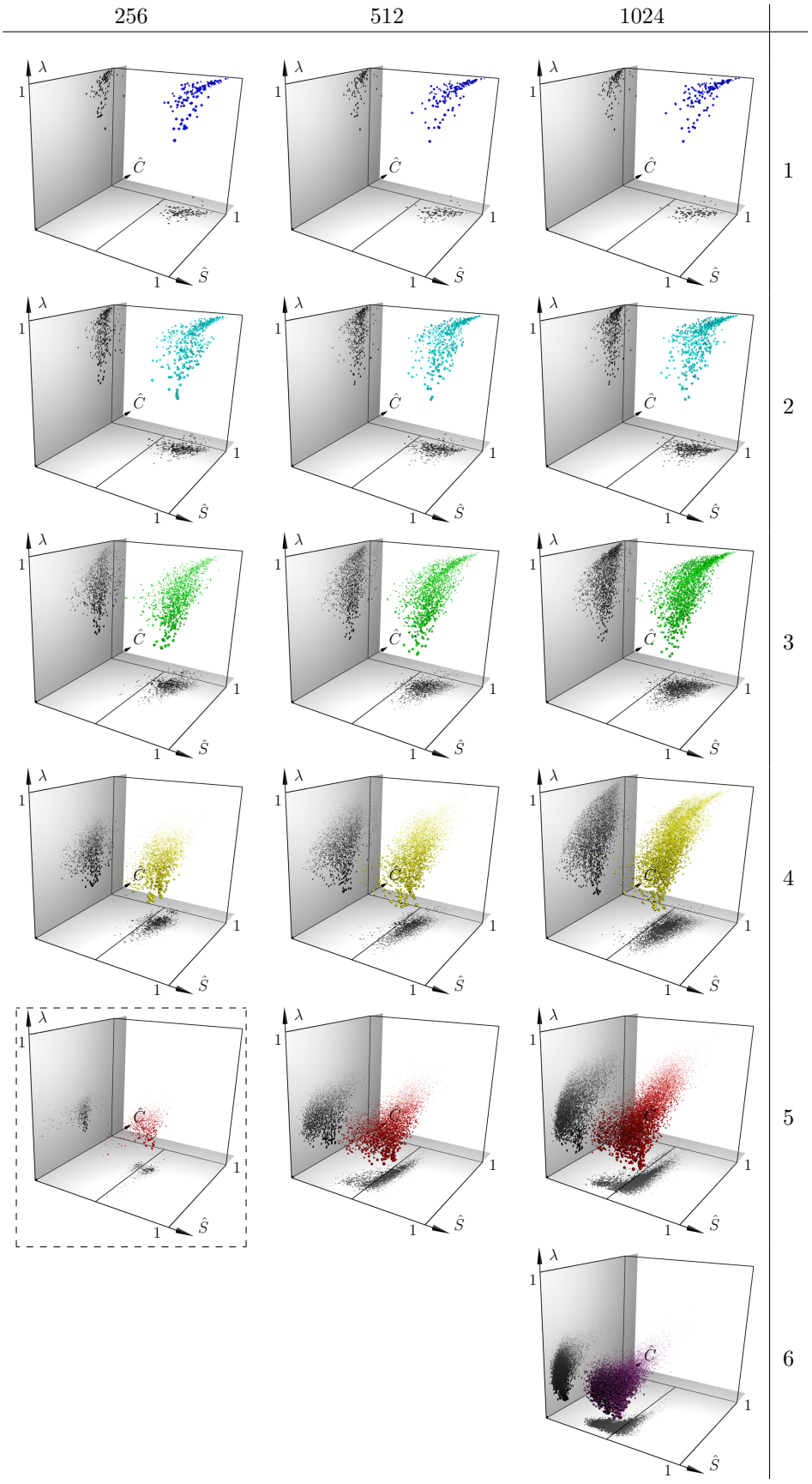


Figure 5.11: Breakdown, by scale number (increasing top to bottom), of 3D views of the visualization spaces for $S_{ij}S_{ij}$ at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

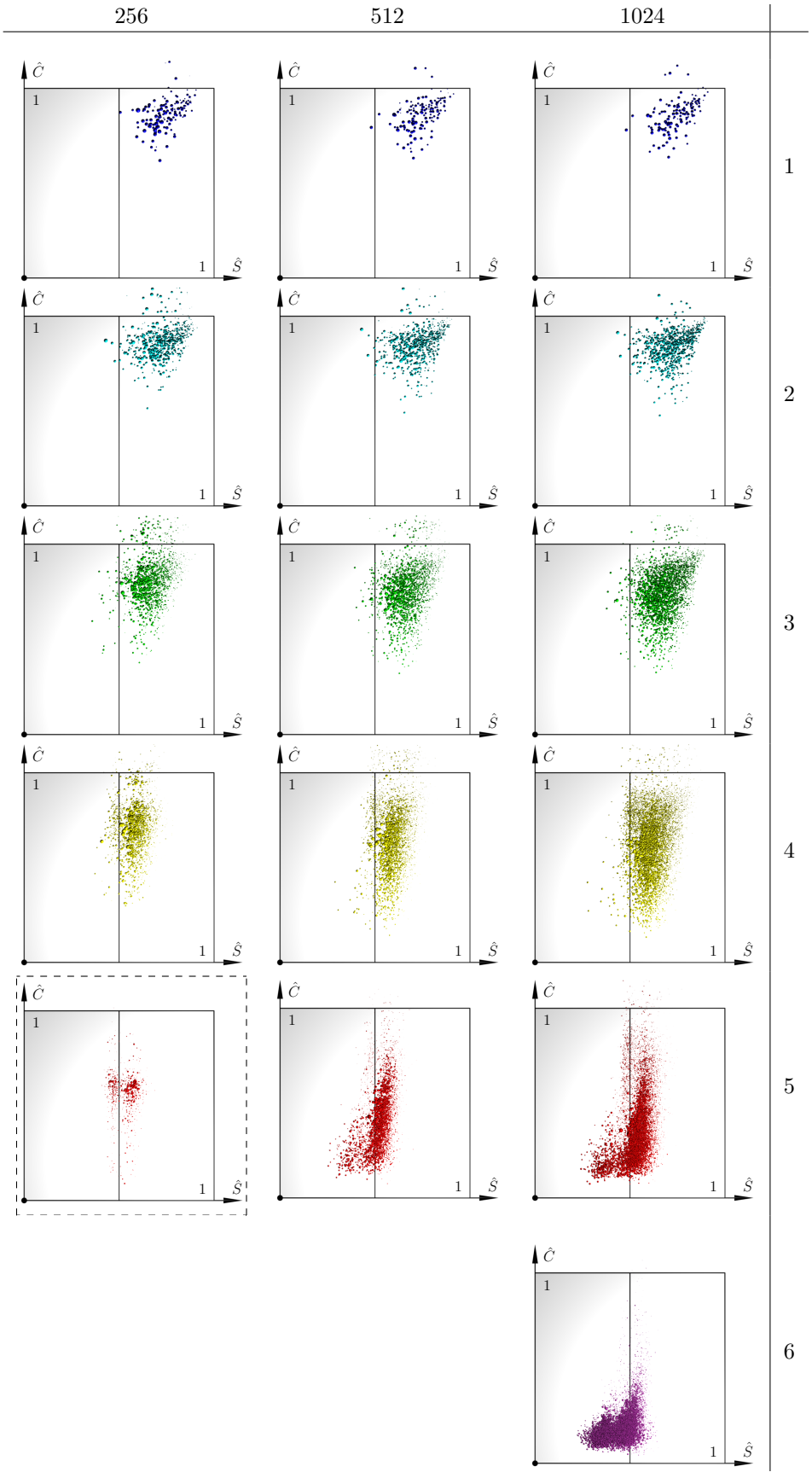


Figure 5.12: Breakdown, by scale number (increasing top to bottom), of top views ($(\hat{S}, \hat{C})$-plane) of the visualization spaces for $S_{ij}S_{ij}$ at 256^3 (left), 512^3 (center), and 1024^3 (right) grid resolutions

tube-like region (see scale numbers 3 and 4 in Figure 5.10), highly stretched particularly for the smaller scales; the transition to sheet-like structures appears to be significant at scale number 4 and becomes obvious at scale number 5, for which dominant structures span across almost all values of $\hat{C}$. On the other hand, $S_{ij}S_{ij}$ structures concentrate less in the tube-like region (see Figure 5.8 and compare scale numbers 3 and 4 of $S_{ij}S_{ij}$ in Figure 5.10 with those of $\omega_i\omega_i$ in Figure 5.12), while they show, at all intermediate scales, many more structures with smaller values of $\hat{C}$, characteristic of sheet-like geometries. The transition to sheet-like structures begins earlier, at scale number 3, for $S_{ij}S_{ij}$ than for $\omega_i\omega_i$, and is completed by scale number 5.

5.4 Effect of grid resolution in the geometry of structures

From Figures 5.7 and 5.8 it is observed that the 256^3 case does not capture well the dominance of sheet-like structures that occurs in both $\omega_i\omega_i$ and $S_{ij}S_{ij}$ in the small scales (scales numbers from 4 on). Figures 5.11 and 5.12 show, for $S_{ij}S_{ij}$, a tendency toward sheet-like structures at the smallest scale studied for the 256^3 case, although the smaller values of $\hat{C}$ present in the 512^3 and 1024^3 for the same scale number are not captured in the 256^3 case either. This is even more pronounced in the $\omega_i\omega_i$ field (see, in particular, Figure 5.10), for which the departure from tube-like region toward the sheet-like structures present at higher grid-resolutions is not obvious at all in the 256^3 case.

The 512^3 case performs better than the 256^3 case in describing the geometry of the structures at the scales of study, when each one is compared with its immediately higher grid resolution. For example, the visualization space at scale number 5, the smallest scale analyzed for 512^3, is rather similar to the 1024^3 case (see central and right columns of Figures 5.9–5.12), where structures with geometries transitioning from the tube-like region to the strongly sheet-like region are captured at both resolutions in more similar proportions. The 1024^3 case still shows a higher concentration of sheet-like structures, particularly for the $S_{ij}S_{ij}$ case.

These results are consistent with the observations of HF. They identified multiple modes of the stretched spiral vortex (Lundgren, 1982) in the numerical database and investigated their formation processes. They found that the highest grid resolution $k_{\max}\bar{\eta} \approx 4$ was needed to eliminate the

fragmentation of sheets for a precise capture of the spiral turns (sheet-like) of those structures and for a proper study of the dissipation field. Schumacher et al. (2005) found also the necessity of resolving sub-Kolmogorov scales when studying the very fine structures in scalar mixing, where sheets are also dominant in the scalar dissipation field. Sreenivasan (2004), based on intermittency arguments, proposed a revised grid resolution criterion, based on estimates of the ratio of maximum to average dissipation obtained from measured multi-fractal exponents (see Sreenivasan & Meneveau, 1988) much more stringent than the traditional $k_{\max}\bar{\eta} \approx 1$ criterion.

As previously noted, the largest and smallest scales for each grid resolution were left out of the analysis. The largest scale is dependent on boundary conditions and forcing applied and, therefore, of less interest in this particular study. The smallest scale was not analyzed to avoid interference with the iso-contouring step due to grid resolution effects. But three grid resolutions of the same flow realization provide the opportunity to verify whether that last statement holds. For that reason, Figures 5.9 to 5.12, include, for the 256^3 case (left column), one additional visualization space corresponding to scale number 5 (framed in a dashed-line box). When compared to the homologous scale number for the higher grid resolutions, it can be seen that the strong sheet-like character of the structures is not well captured in the 256^3. A possible explanation is that sheet-like structures at that resolution are more fragmented into smaller structures (part of the original ones). Some will still be sheet-like but their tube-like area coverage increases since the nearly planar area is reduced and the surface is still closed, which results in higher values of $\hat{C}$ directly affecting their location in the visualization space. Some others can even result in small blob-like structures or, in general, rather distinct geometries than the original sheets of which they are fragments. This is confirmed when the population of individual structures at that scale number is evaluated, since there is a high increase of small-area structures. The pixelization effect seen in Figure 5.6 for the two-dimensional plane cuts, when extended to three-dimensions (where the iso-contours, and thus the individual structures, are obtained) can help to visualize the scenario described above.

5.5 Clustering results for the 1024^3 case

As part of the classification step of the methodology, clustering techniques are applied to the structures obtained from all merged scales under study. This is done independently for structures of $\omega_i\omega_i$ and $S_{ij}S_{ij}$. Only the 1024^3 database is considered here. The number of structures present at each scale largely increases with the scale number (i.e., for smaller and smaller scales). Therefore, geometries of structures of larger scales could be under-represented in the clustering process. To avoid that situation a stratified random sampling with a disproportionate allocation, based on the standard deviation of the population of each scale number, is applied among the present scales prior to the clustering algorithm. See Appendix H for more details.

The results of the clustering algorithm applied individually to $\omega_i\omega_i$ and $S_{ij}S_{ij}$ structures are presented in Figure 5.13. An optimum number of 3 clusters was automatically obtained by the algorithm for $\omega_i\omega_i$, while the structures of $S_{ij}S_{ij}$ were optimally clustered in 2 groups. Optimality scores obtained during the automatic determination of the number of clusters for $\omega_i\omega_i$ (left) and $S_{ij}S_{ij}$ (right) are plotted in Figure 5.14. The optimality score is computed as the mean value of the silhouette coefficient of all the clustered elements minus the standard deviation of those silhouette coefficients. Higher mean values of the silhouette coefficient imply that the elements were clustered in groups where they have a high degree of membership. Small values of the standard deviation of the silhouette coefficients indicate homogeneity in that level of membership among the clustered elements. The combination of a high mean and a low standard deviation is sought for an optimum clustering result. We note that the optimality scores obtained are rather low (less than 0.5, 1 being the maximum achievable), even for the optimum number of clusters. Also, the optimality score associated with the optimum number of clusters determined for each case does not differ significantly from the rest. These two facts are an indication that the elements to cluster are organized as a cloud of points continuously distributed throughout the feature space of parameters used for clustering, instead of being organized in well distinguishable groups that would result in higher optimality scores and more variation among those scores for non-optimal number of clusters. A projection of that feature space is the visualization space where the results have been plotted,

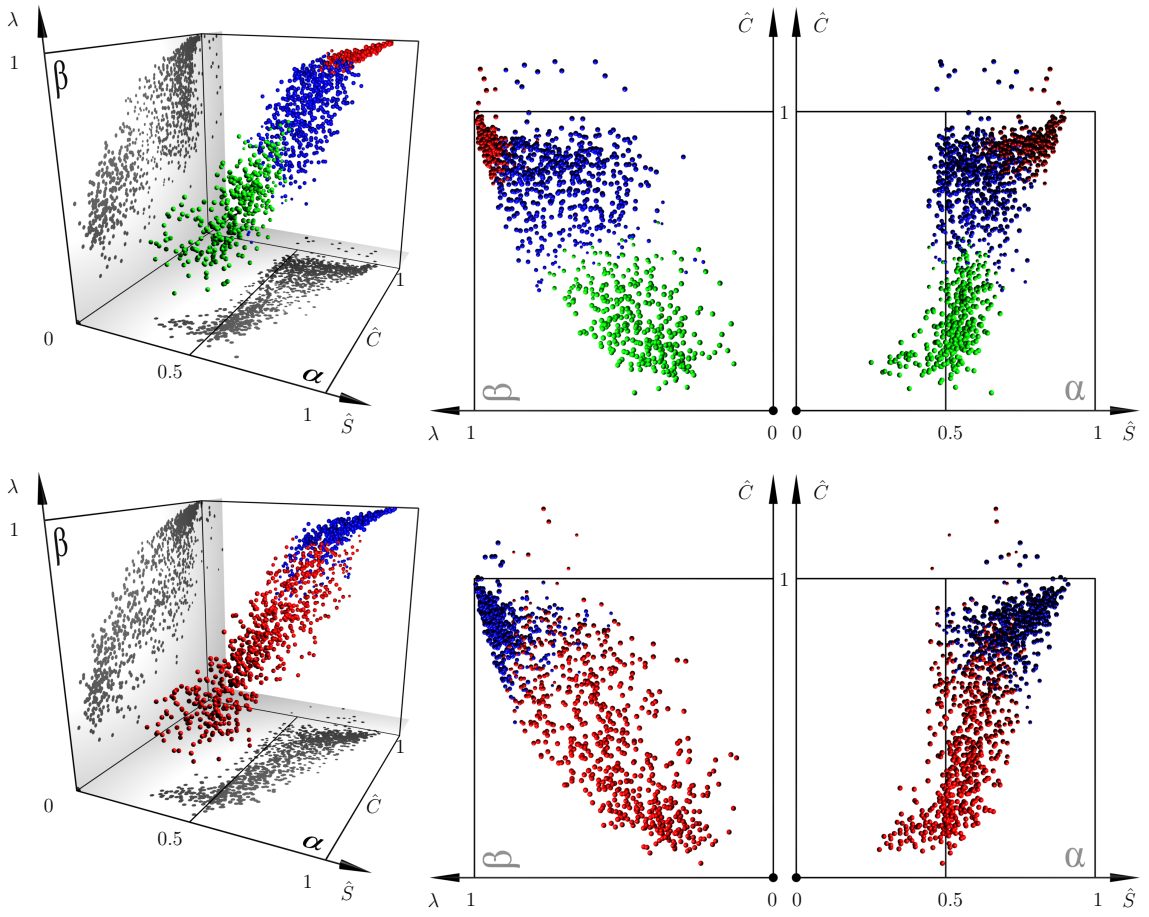


Figure 5.13: Clustering results in the visualization space—3D view (left) and lateral (center) and top (right) projections—with glyphs (spheres) representing the optimum clusters of structures educed from the stratified random sample with optimum allocation of the sets of $\omega_i\omega_i$ (top) and $S_{ij}S_{ij}$ (bottom) structures. Glyphs are scaled by the normalized value of the silhouette coefficient, which indicates the degree of membership of that element to the assigned cluster

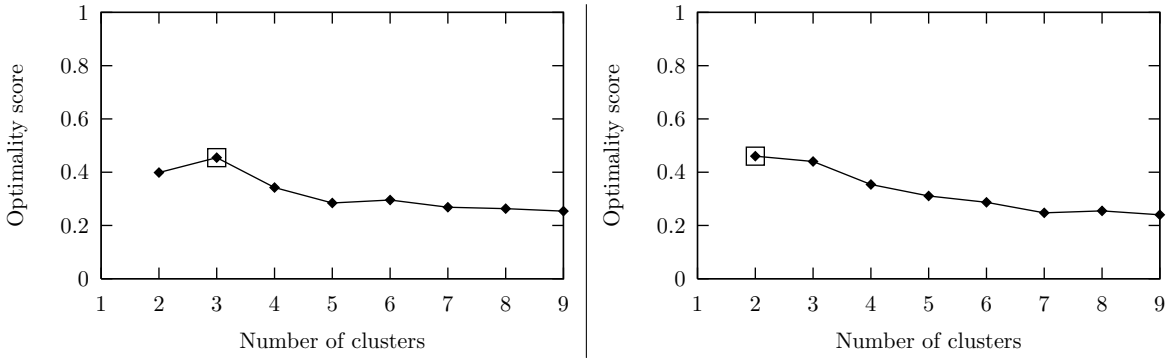


Figure 5.14: Optimality scores for different number of clusters obtained during the application of the clustering algorithm to the set of structures of $\omega_i\omega_i$ (left) and $S_{ij}S_{ij}$ (right) independently. Optimum number of clusters (square point) of 3 and 2 were automatically determined for $\omega_i\omega_i$ and $S_{ij}S_{ij}$, respectively

where it is possible to see also the continuously distributed cloud of glyphs.

5.6 Discussion

The dominance of tube-like structures at intermediate scales of $\omega_i\omega_i$ is consistent with the presence of so-called 'worms' reported in the fluid mechanics literature (see, for example, Siggia, 1981; Jiménez et al., 1993). Tube-like structures appear also at intermediate scales of $S_{ij}S_{ij}$ but in less proportion than for $\omega_i\omega_i$. At all scales analyzed, $S_{ij}S_{ij}$ shows, on average, more planar geometries than $\omega_i\omega_i$. Also the transition to sheet-like structures occurs earlier (larger scale) for $S_{ij}S_{ij}$.

The maximum departure between the spectra of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ occurs at the intermediate scales (scale numbers 3 and 4), as observed in Figure 5.3. This seems to translate into differences in the geometrical character of structures of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ at those scales numbers. In physical space, the higher concentration of tube-like structures found in the enstrophy field might be one geometrical link to its higher intermittency, when compared to the dissipation field.

At the smallest scale, both fields show a clear dominance of sheet-like structures. They appear highly stretched, that is, with small thickness, but their spatial extent can be significant. Instabilities of vortex sheets have been suggested as a primary mechanism responsible for the generation of vortex tubes in turbulent flows. Vincent & Meneguzzi (1994) found that the production of vortex sheets and their subsequent roll-up, forming tubes, shows a strong correlation between scales, and occurs in a one-step process (in contrast with Richardson multi-step cascade picture). Furthermore, they explain the alignment of vorticity with the intermediate strain-rate eigenvector as a consequence of vorticity sheet production by strong strain, instead of tube formation. HF identified the stretched spiral vortex (Lundgren, 1982) in homogeneous isotropic turbulence, appearing in three modes (two symmetric and one antisymmetric), that involve one or multiple vortex sheets interacting to generate tubes. While the stretched spiral vortex was not found in the work of Vincent & Meneguzzi (1994), that was attributed by HF to the increased grid resolution required to avoid fragmentation of the spiral turns. The correlation of geometries for $\omega_i\omega_i$ and $S_{ij}S_{ij}$ at the smallest scale is consistent with the known feature of sheets, in which strain and rotation rates are both large and correlated

(Ruetsch & Maxey, 1992; Horiuti & Takagi, 2005).

Previous studies of the enstrophy field also suggest (see Nomura & Post, 1998, and the references therein) that its geometry depends on its local magnitude: intense regions appear tube-like while moderate enstrophy regions seem to be more sheet-like. These considerations are, nevertheless, independent of the scale. Besides the iso-contour value of the mean plus twice the standard deviation of each filtered component field, whose results have been presented here, we examined also the mean plus three times the standard deviation, with no significant differences of the geometries educed for each field. A wider range of iso-contour values would help clarify the sensitivity of the geometries Nonetheless, we note that the categorization of globally intense or moderate values of the original field (containing all scales) based on the iso-contour value is not directly applicable to its component fields after the multi-scale decomposition.

The 256^3 grid resolution ($k_{\max}\bar{\eta} \approx 1$) was unable to reproduce the predominance of highly stretched sheet-like structures found for the smaller scales at higher grid resolutions. This indicates that sub-Kolmogorov scales must be resolved for a proper geometrical study of the smallest structures in turbulence, as has been suggested in the literature (see Shumacher & Sreenivasan, 2005; Horiuti & Fujisawa, 2008) when studying intermittent fields: their high fluctuations (manifested in the long tails of the volume pdfs in physical space) will occur at very fine scales. These, in general, can be substantially smaller than the average Kolmogorov length scale, $\bar{\eta} = \left(\nu^3/\langle\epsilon\rangle\right)^{1/4}$, defined in terms of the average rate of kinetic energy dissipation per unit mass, $\langle\epsilon\rangle$, and the viscosity of the fluid, ν, and traditionally used to define the largest dynamically significant wavenumber resolved in DNS, $k_{\max}$, such that $k_{\max}\bar{\eta} \approx 1$. As noted in §5.2 the volume pdfs of the different scale component fields obtained from $\omega_i\omega_i$ and $S_{ij}S_{ij}$ show wider ranges for smaller scales, indicating that higher fluctuations of those two fields occur in the small scales and confirming the intermittency of those two fields.

Chapter 6

Assessment of the new non-local methodology complementing existing local methods

In this chapter, the non-local methodology for the study of the geometry of structures in turbulence introduced in Chapter 2 and used in Chapter 5 is utilized to complement two local criteria present in the literature for the identification of vortex tubes and vortex sheets in turbulent flows. The purpose is to provide a qualitative and quantitative assessment of the geometrical aspects of those local identification criteria. This is performed by applying the non-local methodology to confirm whether the educed structures show the expected geometries. The local and non-local methods are applied to the same numerical database referred to in Chapter 5.

6.1 Local identification criteria

Among the various local criteria found in the literature, here we consider those used in Horiuti & Takagi (2005) and HF for educing vortex tubes and sheets, which are outlined below. A point is considered to belong to a vortex tube core where the second invariant, Q, of the velocity gradient tensor, $\partial u_i/\partial x_j$, has a *sufficiently* large value. Q is defined as:

$$Q \equiv \frac{1}{2}\left[\left(\frac{\partial u_i}{\partial x_i}\right)^2 - \frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\right]. \tag{6.1}$$

For incompressible flow, $\partial u_i/\partial x_i = 0$, and Q is related to $\omega_i\omega_i$ and $S_{ij}S_{ij}$ by:

$$2\,Q = -\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i} = \Omega_{ij}\Omega_{ij} - S_{ij}S_{ij} = \frac{1}{2}\,\omega_i\omega_i - S_{ij}S_{ij} \tag{6.2}$$

The condition $Q > 0$ was first used by Hunt et al. (1988), in combination with the additional constraint of the pressure, p, being lower than ambient, to define vortex tubes. The Poisson's equation for pressure in incompressible turbulent flow with density ρ can be rewritten as (see Bradshaw & Koh, 1981):

$$\frac{1}{\rho}\frac{\partial^2 p}{\partial x_i \partial x_i} = -\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i} = 2\,Q = \frac{1}{2}\,\omega_i\omega_i - S_{ij}S_{ij} \tag{6.3}$$

Therefore, Q is a pressure source term. Also, from the latter equality, enstrophy acts as a source term while dissipation acts as a sink term of pressure. There are some situations in which the Q-criterion is not adequate for educing vortex tubes, as is shown in Jeong & Hussain (1995) (e.g., conically symmetric vortex, axisymmetric axial vortex within a vortex ring, Bödewadt vortex).

For educing vortex sheets, the method proposed by Horiuti & Takagi (2005) is used in this chapter. At any given point, the eigenvalues of the symmetric second-order tensor $A_{ij} \equiv S_{ik}\Omega_{kj} + S_{jk}\Omega_{ki}$ are reordered as $[A_{ij}]_\omega$, $[A_{ij}]_+$, and $[A_{ij}]_-$. $[A_{ij}]_\omega$ is the eigenvalue whose corresponding eigenvector is most aligned with the vorticity field, ω_i, at that point. $[A_{ij}]_+$ and $[A_{ij}]_-$ are the remaining largest and smallest eigenvalues (in an algebraic sense), respectively. The eigenvalues, ϑ, of A_{ij} can be obtained from the depressed cubic equation:

$$\vartheta^3 - \frac{1}{2}A_{ij}A_{ji}\,\vartheta + \frac{1}{3}A_{ij}A_{jk}A_{ki} = 0. \tag{6.4}$$

Note that there is no term in ϑ^2 for being $A_{ii} = 0$, due to the symmetry of S_{ij} and antisymmetry of Ω_{ij}. Iso-contours of $[A_{ij}]_+$ are considered vortex sheets. This method takes advantage of the known feature of vortex sheets in which strain rate and vorticity are both large and correlated, reflected in $[A_{ij}]_+$. Horiuti & Takagi (2005) explain the advantages of this identification criterion over previously existing ones also based on that feature of vortex sheets (see, for example, Tanaka & Kida, 1993).

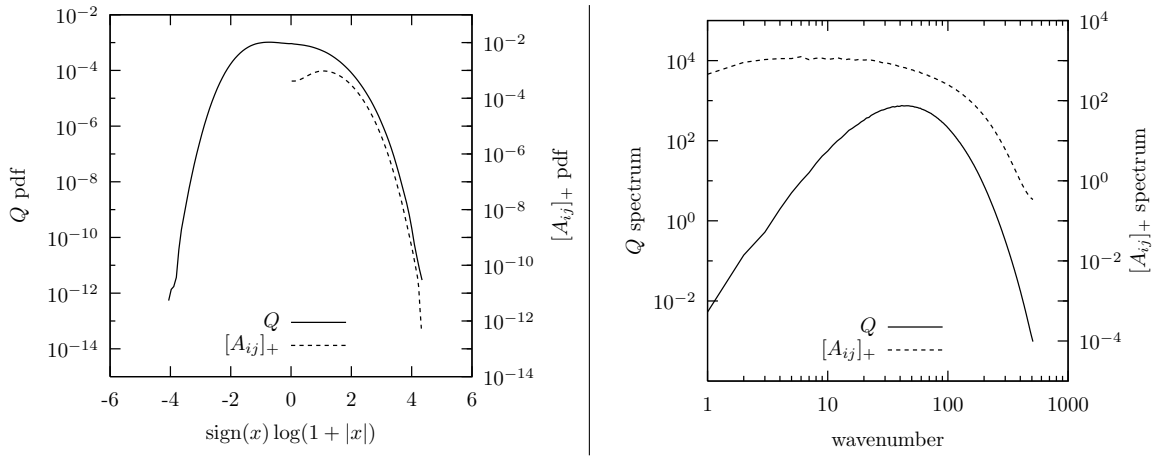


Figure 6.1: Volume pdfs in physical domain (left) and spectra in Fourier domain (right) of Q and $[A_{ij}]_+$ fields for the 1024^3 grid resolution. Note that the volume pdfs use a transformation of the form $\mathrm{sign}(x)\log(1+|x|)$ in the abscissa coordinate, and that curves for Q and $[A_{ij}]_+$ fields use two different vertical axes (both in the pdfs and the spectra), shifted one decade for a clear view (non-intersecting curves)

6.2 Application of non-local methodology

Once the sets of iso-surfaces of Q and $[A_{ij}]_+$ are obtained, the non-local methodology introduced in Chapter 2 is applied to both sets. The multi-scale decomposition is not used in the extraction step, since the purpose is to assess the geometrical character of the iso-surfaces extracted by the local criteria. A multi-scale decomposition of the Q and $[A_{ij}]_+$ scalar fields could be applied beforehand (as was done for $\omega_i\omega_i$ and $S_{ij}S_{ij}$ in Chapter 5), and then iso-contours of the component fields could be independently obtained, but the meaning of the educed structures would not be the same as those obtained by iso-contouring the original fields of Q and $[A_{ij}]_+$, and the purpose of the assessment of the local criteria would be lost. While the three grid resolutions are available also for Q and $[A_{ij}]_+$ fields, only the finest (i.e., 1024^3) is used, since this chapter is not intended to evaluate the effect of the grid resolution in Q and $[A_{ij}]_+$ structures.

Figure 6.1 shows plots of the volume pdfs (left) and the spectra (right), in Fourier space, of the two fields, Q and $[A_{ij}]_+$, for that finest grid resolution (1024^3). The two plane cuts in Figure 6.2 correspond to Q (left) and $[A_{ij}]_+$ (right) fields, and were obtained normally to one of the principal directions of the cubic domain at half the side length. Iso-surfaces of Q and $[A_{ij}]_+$ extracted at contour values equal to the mean plus 5 and 4 times, respectively, the standard deviation of each

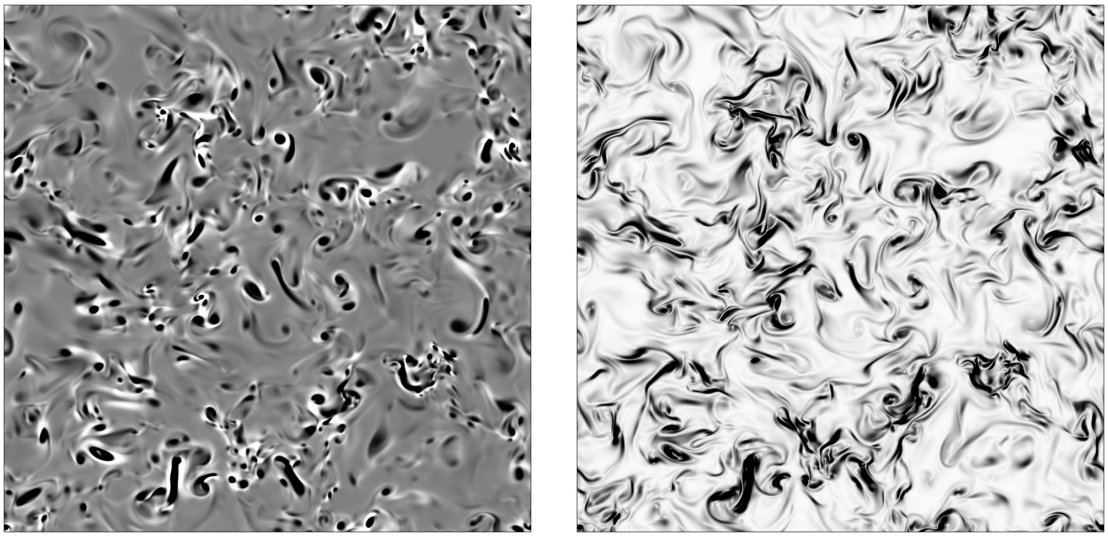


Figure 6.2: Plane cuts of Q (left) and $[A_{ij}]_+$ (right) fields normal one of the principal directions of the cubic domain, at half its side length

field are presented in Figure 6.3. These contour values are approximately the same ones used in HF (1200 for Q and 1000 for $[A_{ij}]_+$) to educe vortex tubes and sheets. The visualization spaces in Figure 6.4 contain glyphs corresponding to the geometrical characterization of the individual structures shown in Figure 6.3, with the same coloring scheme (blue used for Q structures and red for $[A_{ij}]_+$ structures). It is observed that structures of Q tend to be located near the $(\hat{S}, \hat{C}) = (1/2, 1)$ region, where tube-like structures are generally located, and present moderate-to-high stretching. On the contrary, $[A_{ij}]_+$ structures appear much closer to the $\hat{C} = 0$ region, thus corresponding to more planar geometries (sheet-like), and with lower values of λ, implying more stretched structures.

The clustering algorithm described in the classification step of the non-local methodology is applied to the set of structures formed by the union of both sets of iso-surfaces of Q and $[A_{ij}]_+$ fields, without any a priori distinction of structures of those two sets. This means that the algorithm has no knowledge of whether individual structures were extracted from Q or $[A_{ij}]_+$ fields. Owing to the equivalent number of structures educed for Q and $[A_{ij}]_+$ and the similar standard deviations of each population, the pre-clustering stratified random sampling with disproportionate allocation in this case results in practically the union of the complete sets. Then, structures are clustered based solely on their geometrical characterization given by set of parameters $\{\hat{S}, \hat{C}, \lambda, d_u^S, d_l^S, d_u^C, d_l^C\}$, that

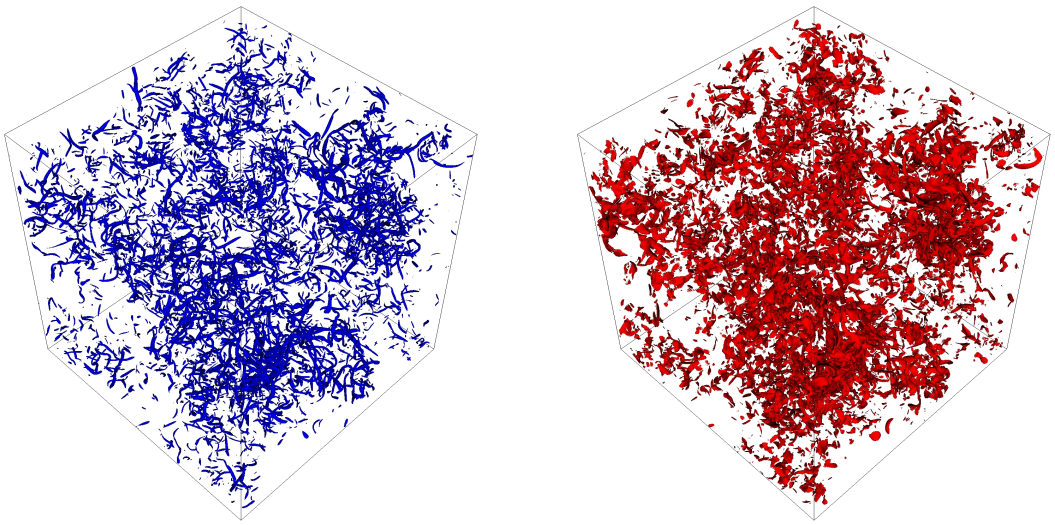


Figure 6.3: Iso-contours of Q and $[A_{ij}]_+$ fields extracted at their mean plus 5 and 4 times their standard deviation, respectively

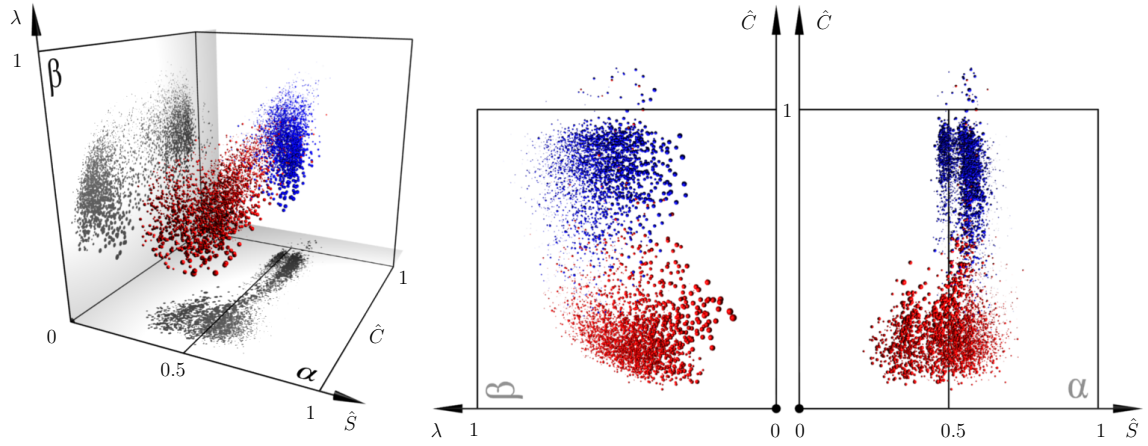


Figure 6.4: Visualization space—3D view (left) and lateral (center) and top (right) projections—with glyphs (spheres) representing educed structures of Q (blue) and $[A_{ij}]_+$ (red), scaled by the lognormalized area of each corresponding structure

define the feature space described in §2.3.2. An optimum number of clusters of 2 was found (see Figure 6.6), and the resulting clusters are shown in Figure 6.5, where each color corresponds to a different cluster. Glyphs (spheres) in that figure are scaled by the normalized silhouette coefficient, a measure of the degree of membership of each structure to the cluster it was assigned. Comparing Figures 6.5 and 6.4, it is confirmed that the two educed clusters correspond in their majority to the two sets of structures (Q and $[A_{ij}]_+$). Numerically a matching of 96% between pairs of groups was

obtained. A small percentage of structures of both fields shows a different geometry than the one expected according to the local criterion. For example, the central tube in Figure 7.5 (discussed later) is a structure of $[A_{ij}]_+$, which is meant to educe sheet-like structures, while the other two tubes in that plot are structures of Q, as expected.

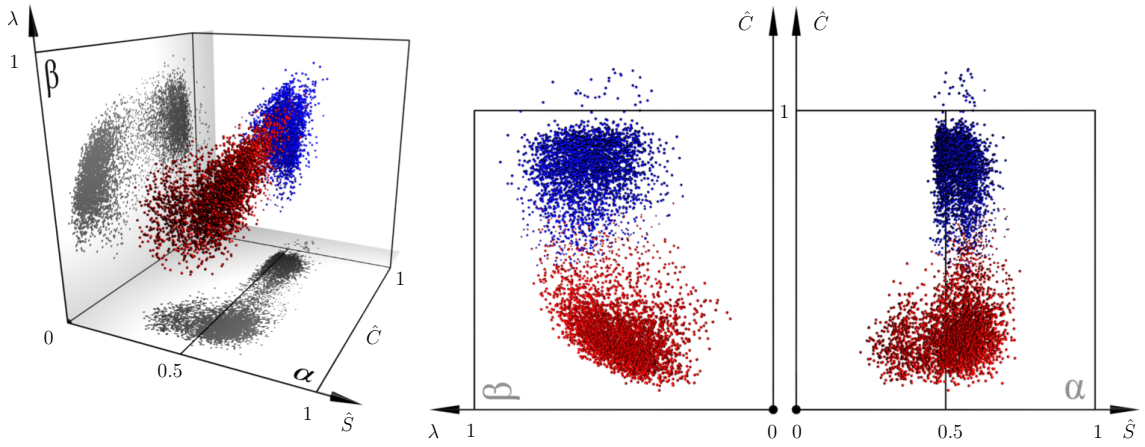


Figure 6.5: Clustering results in the visualization space—3D view (left) and lateral (center) and top (right) projections—with glyphs (spheres) representing optimum clusters of structures educed from the set union of Q and $[A_{ij}]_+$ structures. Glyphs are scaled by the normalized silhouette coefficient, which indicates the degree of membership of that element to the assigned cluster

We emphasize that the scaling of the glyphs in Figures 6.4 and 6.5 is different. Figure 6.4 uses the lognormalized area of each structure, and thus it can be concluded, for both Q and $[A_{ij}]_+$ that more stretched structures are typically larger (in area): glyphs with lower values of λ (more stretched) are bigger in that figure. On the other hand, glyphs in Figure 6.5 are scaled by the silhouette coefficient of the corresponding structure as a result of the clustering algorithm: bigger glyphs imply larger silhouette coefficient and, therefore, a higher degree of membership to the cluster the structure was automatically assigned. This different scaling contributes to an apparent lower density of glyphs in Figure 6.4, when compared to Figure 6.5, since glyphs associated with small structures would be also small in the former, while they will appear larger in the latter, if they have a high silhouette coefficient.

Figure 6.6 shows the optimality scores found during the automatic determination of the number of clusters. It is observed that a number of clusters of 2 provides an optimality score near 0.7 (1.0 being

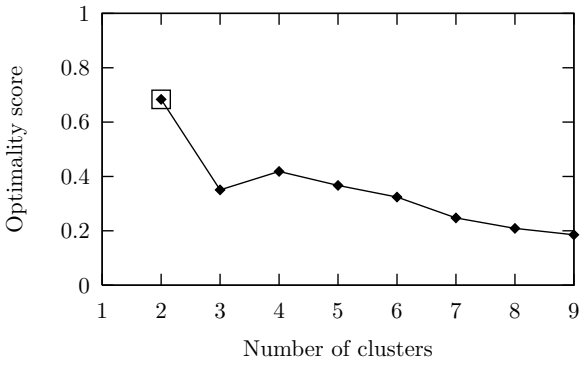


Figure 6.6: Optimality scores for different number of clusters obtained during the application of the clustering algorithm to the set of structures of Q and $[A_{ij}]_+$ together. An optimum number of clusters of 2 was automatically determined (square point)

the maximum), much higher, comparatively, than for other numbers of clusters. A comparison with the clustering results obtained for $\omega_i\omega_i$ and $S_{ij}S_{ij}$ in §5.5 (see Figure 5.14), where the optimality scores were rather low and the variation among different number of clusters was small, gives an indication of the higher level of confidence in the clustering results for this case. This could be anticipated by looking at the organization of glyphs in the visualization space in Figure 6.5, when compared to Figure 5.13.

Thus, the intuition that resulted by the visual cues of Figure 6.3 in which iso-surfaces of Q seemed tube-like and structures of $[A_{ij}]_+$ appeared sheet-like has been verified with a mathematical foundation first by the geometrical characterization, whose results can be partially seen in Figure 6.4, and then with the application of clustering techniques to the union of both sets, whose results are seen in Figure 6.5. The geometrical data can be used both qualitative and quantitatively.

Chapter 7

Interaction among structures of different fields: proximity issues

This chapter is devoted to the study of the interplay among structures of different fields under a geometrical perspective. We present here a methodology that, for each individual structure, performs an analysis of its proximal structures based upon their geometrical characterization previously introduced. Statistical results are obtained for the set of structures under study, presented in the form of combined probability density functions.

A motivation for this methodology is presented in §7.1, followed by the description of the methodology itself in §7.2. Results of its application to the four fields studied in previous chapters of this thesis grouped by pairs, namely Q–$[A_{ij}]_+$ and $\omega_i\omega_i$–$S_{ij}S_{ij}$, are presented in §7.3 and §7.4, respectively.

7.1 Motivation

The scalar fields that have been studied in the previous two chapters were all derived from the velocity gradient tensor. Some of them can be formulated in terms of the others by simple algebraic relations (e.g., Q in terms of $\omega_i\omega_i$ and $S_{ij}S_{ij}$). Other fields, not studied here but common in the study of turbulence, such as the pressure field, p, are also related to these (see equation 6.3). A passive scalar field could be also added to the flow and related to the other fields through the advection-diffusion equation, as we did in Chapter 4. They all can be thought of as different manifestations

of the same flow, in this case, incompressible homogeneous isotropic turbulence decaying in time in a periodic box. The mathematical relations among all of them are well known, and a vast effort has been dedicated in the fluid mechanics literature to study how those mathematical relations are translated into the physical aspects of turbulence, both in physical and Fourier domains, through the study of pdfs, structure and (auto-)correlation functions, spectra, etc.

It is thus conceivable that the structures extracted from those fields may have some relations, first, in their relative locations in physical space and, second, in their geometrical character, forming composite structures localized in physical domain. Perhaps the most common example of such inter-relation between scalar fields is the formation process of a vortex tube that results from the roll-up of a vortex sheet (see HF and the references therein): at an intermediate stage of that process, the core of the vortex tube is dominated by high values of vorticity, while the sheet that is rolling up around it presents high values of dissipation. Therefore, in that scenario, tube-like structures of $\omega_i\omega_i$ would be surrounded by sheet-like structures of $S_{ij}S_{ij}$. Similarly, considering the scalar fields used by the local identification criteria in §6.1 to educe vortex tubes and sheets, structures of Q (which were found to be tube-like) would be surrounded by structures of $[A_{ij}]_+$ (predominantly sheet-like).

Plane cuts of pairs of scalar fields, taken at the same location, are superimposed in Figure 7.1 ($S_{ij}S_{ij}$ over $\omega_i\omega_i$ on the left, $[A_{ij}]_+$ over Q on the right). Close relations between their corresponding structures are noticed: structures of $S_{ij}S_{ij}$ and $[A_{ij}]_+$ (red) tend to wrap around those of $\omega_i\omega_i$ and Q (blue), respectively. It is also observed that many structures of $\omega_i\omega_i$ and Q appear to have circular cores, while structures of $S_{ij}S_{ij}$ and $[A_{ij}]_+$ are more elongated. When extrapolated to the three-dimensional fields, those circular patches of the plane cuts of $\omega_i\omega_i$ and Q will likely belong to tubes, while the elongated regions of plane cuts of $S_{ij}S_{ij}$ and $[A_{ij}]_+$ will probably correspond to sheets around them. Another scenario in which circular and elongated regions of the plane cuts correspond, respectively, to blob-like and tube-like structures in three dimensions would be also possible, but its frequency of occurrence is comparatively smaller, as is concluded from the study of the geometry of the four fields previously done in Chapter 5. A methodology enabling study, in three dimensions, of the geometry of structures of different fields surrounding those of a particular field would be useful

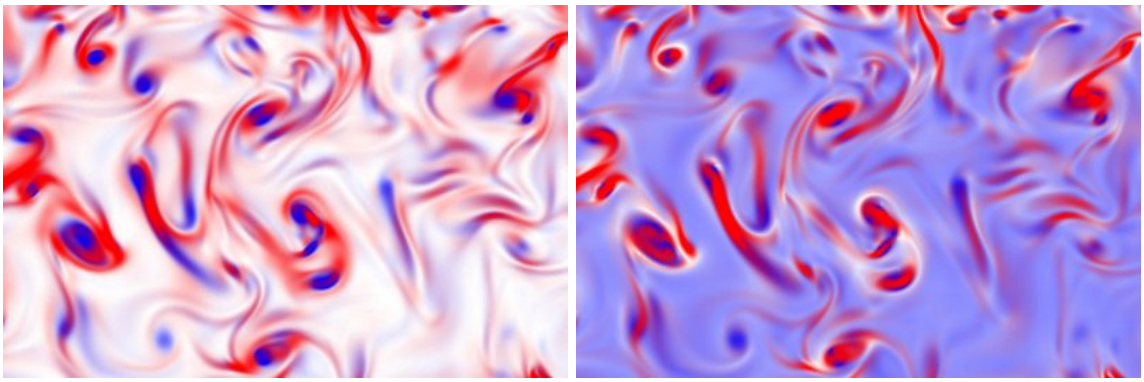


Figure 7.1: Left: Plane cut of $S_{ij}S_{ij}$ (red) superimposed over equivalent plane cut of $\omega_i\omega_i$ (blue). Right: Plane cut of $[A_{ij}]_+$ (red) superimposed over equivalent plane cut of Q (blue)

to test this visual intuition and to quantify its appearance.

7.2 Methodology

Consider two sets of structures, $\mathcal{A}$ and $\mathcal{B}$, containing $N_\mathcal{A}$ and $N_\mathcal{B}$ elements, respectively. We impose no Boolean restriction on both sets, so that elements of $\mathcal{A}$ can also be elements of $\mathcal{B}$.

7.2.1 Processing individual structures

For each structure $a_i \in \mathcal{A}$:

1. Obtain the subset $\mathcal{C}$ of $N_\mathcal{C}$ structures of $\mathcal{B}$ ($\mathcal{C} \subseteq \mathcal{B}$) that are closest to a_i in the bounding box sense. This step is intended to speed up the algorithm by reducing the load required to perform the rest of the steps.

2. For each $c_j \in \mathcal{C}$, $j = 1, ..., N_\mathcal{C}$, obtain the point-wise minimum distance map from c_j to a_i, MDM(c_j, a_i). This map computes, for each point of the discretized surface c_j, the minimum of the distances from it to all points of a_i.

   At the same time, during the computation of MDM(c_j, a_i), those points of a_i that provide the minima for MDM(c_j, a_i) are tagged and an array of N_γ parameters is stored, for each one of those points, in a point-wise conditional array map of a_i, CAM(a_i). The array of parameters $\{\gamma_k, k = 1, ..., N_\gamma\}$ contains, for each tagged point $P \in a_i$, the (dimensionless) proximity value,

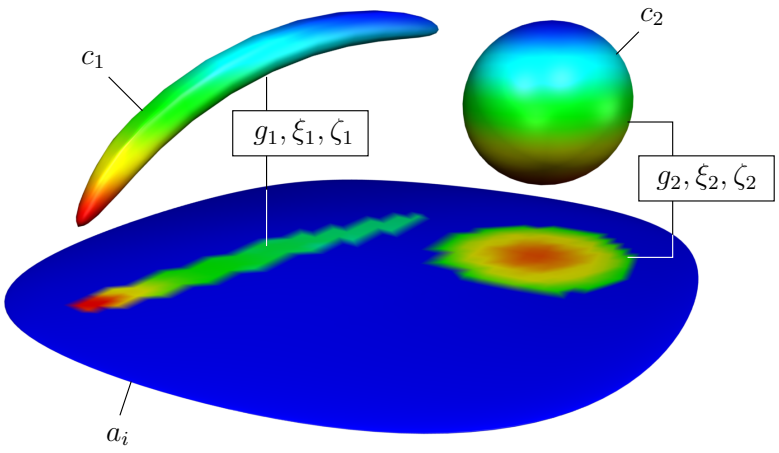


Figure 7.2: Schematic of the computation of the conditional array map (CAM) for a given structure a_i (sheet-like) with respect to two close structures c_1 (tube-like) and c_2 (blob-like). Minimum distance maps from c_j $(j = 1, 2)$ to a_i are shown, with a common color scale for both ranging from red to blue, representing smaller to larger minimum distances to a_i, respectively. The proximity value derived from them is mapped on a_i, with a color scale varying from blue (nil proximity) to red (maximum proximity). Those points of a_i tagged during the computation of the minimum distance maps (with proximity values greater than zero) will store also the values $\{g_j, \xi_j, \zeta_j\}$ of the corresponding c_j in the conditional array map

p, obtained from the minimum distance, d, to the point(s) of c_j for which P was the closest of all points in a_i, plus additional information of c_j itself (for example, but not limited to, an identifier of the structure c_j, and geometrical parameters of c_j obtained from a previous characterization performed on it). The dimensionless proximity, p, is defined in terms of the distance d by $p \equiv (1 + d/D_{a_i})^{-1}$, where D_{a_i} is a non-dimensionalizing length scale of a_i. Smaller distances, d, translate into higher proximity values ($p \in [0, 1]$).

We consider here the particular case in which, for each point P of a_i, the array $\{\gamma_k, k = 1, ..., N_\gamma\}$ consists of four parameters ($N_\gamma = 4$), redefined for simplicity as $\{p, g, \xi, \zeta\}_P$: p is the dimensionless proximity value to c_j described above, g is an index categorizing c_j among N_g known groups of structures present in $\mathcal{B}$, $\{G_g \subset \mathcal{B}, g = 1, ..., N_g \| \bigcup_{g=1}^{N_g} G_g = \mathcal{B}\}$ and ξ and ζ are geometrical properties of c_j. A schematic of the computation of the conditional array map for a_i based on the minimum distance maps for c_j is depicted in Figure 7.2.

As different c_j are processed, the conditional array map for a_i, CAM(a_i), is updated at previ-

ously untagged points of a_i. Also, the array of parameters of a previously tagged point of a_i is updated if the new proximity value for that point is larger than the previously stored one, meaning that the corresponding distance of the new structure c_j to that point is smaller than for all the previous c_j structures processed.

If c_j is the same structure as a_i, which can happen, since, as we stated above, $\mathcal{A}$ and $\mathcal{B}$ could share elements, then it is discarded in the computation of CAM(a_i). Otherwise it would eclipse all other structures of $\mathcal{C}$ and the conditional array map would be trivial and useless. Nevertheless, if c_j is not the same as a_i, but they happen to be identical, then c_j is included in the computation of CAM(a_i), resulting in a trivial but useful result. For example, if $\mathcal{A}$ and $\mathcal{B}$ are disjoint but their structures are identical by pairs $\{a_i, b_i\}$, the trivial CAM(a_i) obtained for each a_i after processing all structures $c_j \in \mathcal{C}$ would prove their identicality.

3. After all c_j have been processed for a given a_i we obtain, from the CAM(a_i), the joint probability density function (jpdf), in terms of p and g, based on area-coverage, denoted by $\mathcal{P}(p,g)\,|_{a_i \leftarrow \mathcal{C}}$. The value $\int_{p_1}^{p_2} \mathcal{P}(p,g)\,|_{a_i \leftarrow \mathcal{C}}\, \mathrm{d}p$ can be interpreted as the probability of finding structures of $\mathcal{C}$ categorized in a group G_g as being the closest structures to a_i in the range of proximity values $[p_1, p_2]$. At a given point $P \in a_i$, a structure c_{j^*} is the closest to a_i at P if it has the highest value of proximity, p, among all the structures c_j that would tag P in the computation of CAM(a_i).

   We compute also the area-based joint probability density function in terms of ξ and ζ, with an additional intensity component based on the averaged proximity value. For each 2D interval of the geometrical properties $[\xi_m, \xi_m + \Delta\xi] \times [\zeta_n, \zeta_n + \Delta\zeta]$ the discrete joint probability density function with intensity (jpdf+i) has two components: the first one is the pdf value itself, i.e., surface area of the a_i such that $(\xi, \zeta) \in [\xi_m, \xi_m + \Delta\xi] \times [\zeta_n, \zeta_n + \Delta\zeta]$ divided by the total area of a_i; the second component is the area-weighted average of the proximity values of all faces of the discretized a_i such that $(\xi, \zeta) \in [\xi_m, \xi_m + \Delta\xi] \times [\zeta_n, \zeta_n + \Delta\zeta]$. We denote that jpdf+i by $\vec{\mathcal{PI}}(\xi, \zeta; p)\Big|_{a_i \leftarrow \mathcal{C}}$, where the vector symbol reflects that it has two components (pdf in terms of $\{\xi, \zeta\}$ and intensity based on p).

We approximate:

$$[\mathcal{P}(p,g)|_{a_i \leftarrow \mathcal{B}} \approx [\mathcal{P}(p,g)|_{a_i \leftarrow \mathcal{C}}\,, \qquad \left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{a_i \leftarrow \mathcal{B}} \approx \left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{a_i \leftarrow \mathcal{C}}\,. \tag{7.1}$$

This approximation is exact when the structures of $\mathcal{C}$ eclipse, for $a_i \in \mathcal{A}$, structures of $\mathcal{B}-\mathcal{C}$, if $N_\mathcal{C} < N_\mathcal{B}$, and also in the trivial case $N_\mathcal{C} = N_\mathcal{B}$, for which it becomes an identity.

7.2.2 Transition from individual structures to results for the set $\mathcal{A}$

Once all $a_i \in \mathcal{A}$ have been processed, global jpdf and jpdf+i are obtained for the set $\mathcal{A}$ as the average of individual jpdf and jpdf+i for each a_i:

$$[\mathcal{P}(p,g)|_{\mathcal{A} \leftarrow \mathcal{B}} = \frac{\sum_{a_i \in \mathcal{A}} [\mathcal{P}(p,g)|_{a_i \leftarrow \mathcal{B}}}{N_\mathcal{A}} \tag{7.2}$$

$$\left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{\mathcal{A} \leftarrow \mathcal{B}} = \frac{\sum_{a_i \in \mathcal{A}} \left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{a_i \leftarrow \mathcal{B}}}{N_\mathcal{A}} \tag{7.3}$$

respectively. This is equivalent to assigning a probability density of $[\mathcal{P}(p,g)|_{a_i \leftarrow \mathcal{B}}\,/N_\mathcal{A}$ to each event $a_i \leftarrow \mathcal{B}$ and then computing the probability density of the event $\mathcal{A} \leftarrow \mathcal{B}$ as the union of all individual events ($\forall a_i \in \mathcal{A}$), taking them as independent. The same reasoning applies to $\left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{\mathcal{A} \leftarrow \mathcal{B}}$.

7.2.3 Computational remarks

When dealing with the discretized surfaces representing the structure a_i, $[\mathcal{P}(p,g)|_{a_i \leftarrow \mathcal{B}}\,/N_\mathcal{A}$ and $\left[\vec{\mathcal{PI}}(\xi,\zeta\,;p)\right|_{\mathcal{A} \leftarrow \mathcal{B}}$ are obtained from face-wise data, and not from point-wise data. This requires a transformation from point-wise to face-wise data, which is done in this case by assigning to each face the array $\{\gamma_k, k = 1, ..., N_\gamma\}$ of its vertex with the maximum value of the proximity (i.e., minimum distance). Interpolation, in this case, is inappropriate since there is no guarantee of continuity of the values of the parameters γ_k throughout the surface, as neighboring points can have data corresponding to different surrounding structures $c_j \in \mathcal{C}$. When continuity of the jpdf variables exists, an interpolation (for example, tri-linear if the faces are triangular) is justified and can improve

the accuracy of the jpdf, relaxing the dependence between the discretization of the surface and the discretization of the jpdf; this was the case in the computation of the $\mathcal{P}(S,C)$ used to obtain the signatures of the structures in the characterization step of the study of the non-local geometry of structures, but cannot be applied to the jpdf and jpdf+i we are analyzing in this chapter.

The algorithm can be modified to obtain directly face-wise data instead of point-wise data and avoid such transformation. Point-wise data has been chosen in our implementation due to the simplicity and increased speed when computing distance maps. Nonetheless, we note that this imposes the additional constraint that both discretized surfaces a_i and c_j must have an equivalent grid resolution for an accurate computation of the distance maps, basis of this algorithm. In our implementation, this is guaranteed as a consequence of the iso-contouring algorithm in use and the fact that, even when multi-scale techniques were applied, sub-sampling was not performed on the grid for any scale, and no decimating operation was applied over the discretized structures thereafter.

7.3 Application to structures of Q and $[A_{ij}]_+$

We apply the methodology explained in §7.2 to the structures of the fields Q and $[A_{ij}]_+$ educed in §6.1. In our first application, the set $\mathcal{A}$ will be composed of the extracted structures of Q, $\mathcal{A} \equiv \mathcal{X}(Q)$, where $\mathcal{X}(\alpha)$ denotes the set of extracted structures from a three-dimensional scalar field α, while the set $\mathcal{B}$ is composed of the union of structures extracted from Q and those extracted from $[A_{ij}]_+$, $\mathcal{B} \equiv \mathcal{X}(Q) \cup \mathcal{X}([Aij]_+)$. Therefore, in this particular case, $\mathcal{A} \subset \mathcal{B}$.

We choose (ξ, ζ) as the geometrical parameters $(\hat{S}, \hat{C})$ obtained in the characterization step of the non-local study of the geometry of structures applied in §6.2. The index g refers in this case to the two groups of structures in which $\mathcal{B}$ can be immediately divided, namely, structures of Q ($g = 1$) and structures of $[A_{ij}]_+$ ($g = 2$). The length scale D_{a_i} used to non-dimensionalize distances when computing proximity values for each structure a_i is taken to be the parameter $\mu \equiv 3\,V/A$, where V is the volume and A the area of a_i. Note that this parameter was used in the geometrical characterization step of Chapter 2 to non-dimensionalize the curvedness (see equation 2.6), resulting in the dimensionless value C, from which $\hat{C}$ was obtained. For a sphere of radius R, $\mu_{sphere} = R$;

for an elongated tube with circular cross section of radius R, $\mu_{tube} \approx 3\,R/2$; for a predominantly sheet-like structure of thickness t, $\mu_{sheet} \approx 3\,t/2$. After applying the methodology, we obtain both $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right|_{\mathcal{A}\leftarrow\mathcal{B}}$ and $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$.

7.3.1 Proximity and area coverage of surrounding structures through jpdf+i

Figure 7.3 shows $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right|_{\mathcal{A}\leftarrow\mathcal{B}}$. The top left plot is the representation of the pdf component, using a greyscale, where white indicates nil area coverage and black indicates the maximum area coverage. Therefore, dark regions indicate that structures of $\mathcal{B}$ with values of $(\hat{S},\hat{C})$ within those regions are found, on average, to surround comparatively a larger proportion of the area of structures of $\mathcal{A}$. Light regions indicate values of $(\hat{S},\hat{C})$ not so commonly encountered in the structures of $\mathcal{B}$ surrounding those of $\mathcal{A}$. The top-right plot represents the intensity component, which in this case corresponds to the proximity value, using a color-scale (continuous gradient blue-cyan-green-yellow-red): blue indicates low proximity (farther distances) while red indicates high proximity (closer distances) of the structures of $\mathcal{B}$ with given values $(\hat{S},\hat{C})$ to those of $\mathcal{A}$. The bottom-left plot is a composition of both the pdf component and the intensity component by means of a hue-saturation gradient, represented in the bottom-right plot. In a HSB (hue, saturation, brilliance) color space, the area-coverage corresponds to the saturation, S, while the intensity component corresponds to the hue, H, and the brilliance, B, is kept constant at its maximum value. Therefore, it contains information of the averaged area-coverage in the saturation scale and information of the averaged proximity in the hue scale. Saturated-red regions correspond to values of $(\hat{S},\hat{C})$ found in structures of $\mathcal{B}$ closest to those of $\mathcal{A}$ and covering, comparatively, the largest proportion of their surface area. Desaturated regions indicate less area coverage and colors closer to the blue hue indicate lower values of the proximity (and thus, farther structures). Regions of saturated-blue, for example, will indicate that structures of $\mathcal{B}$ with those values of $(\hat{S},\hat{C})$ appear far but cover a large proportion of the surface area of structures of $\mathcal{A}$. Desaturated blue regions, on the other hand, will indicate that structures of $\mathcal{B}$ with those values of $(\hat{S},\hat{C})$ appear far as well, but covering a small fraction of the surface area

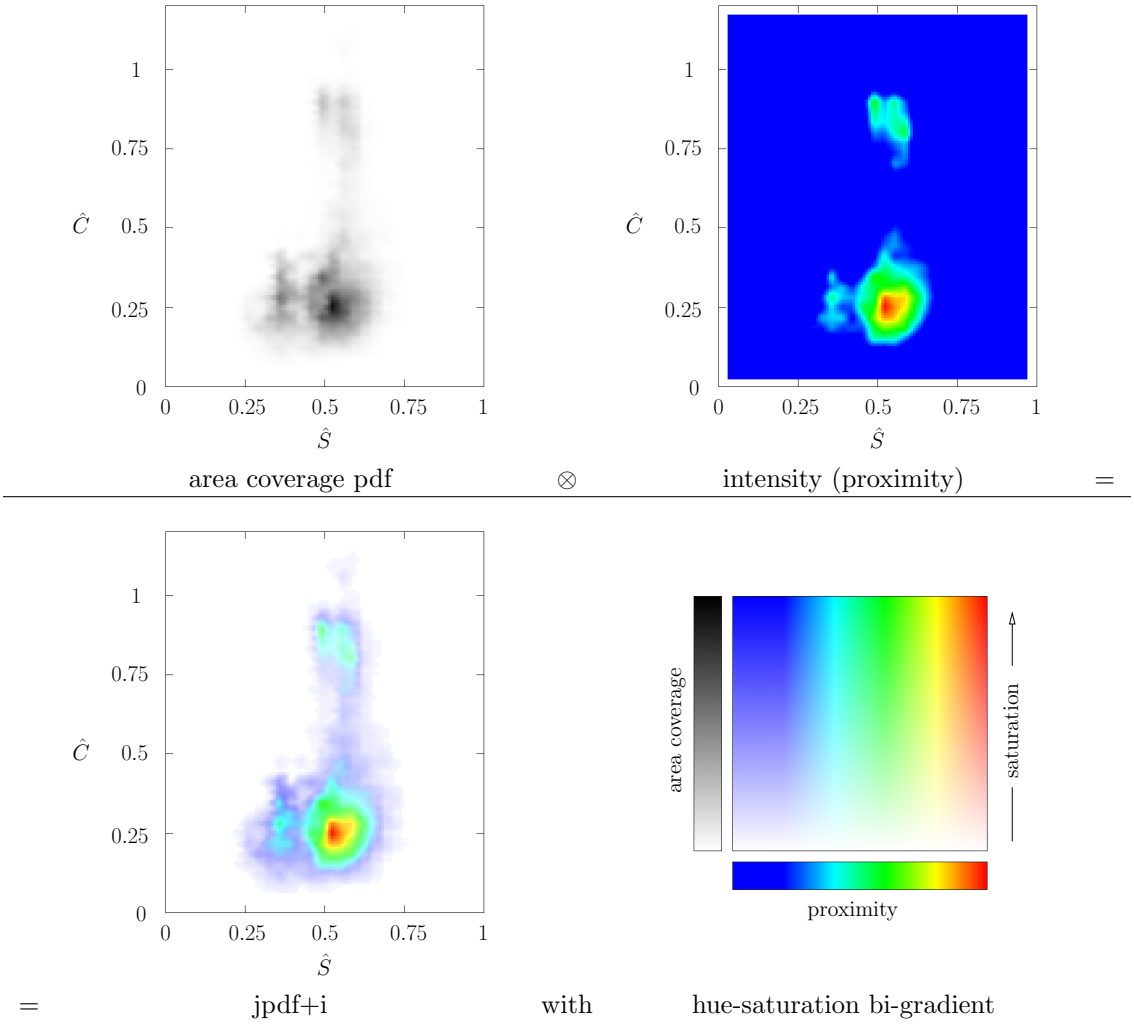


Figure 7.3: Components of the jpdf+i in terms of $(\hat{S}, \hat{C})$, plus intensity component based on proximity, of structures of $\mathcal{X}(Q) \cup \mathcal{X}([A_{ij}]_+)$ surrounding structures of $\mathcal{X}(Q)$: area-coverage pdf component (top left) using greyscale; intensity component (top right) using blue-cyan-green-yellow-red color-scale; composition of area-coverage pdf and intensity components to obtain the composite plot (bottom left) with bi-dimensional hue-saturation gradient scale (bottom right) corresponding to proximity (hue) and area coverage (saturation)

of structures of $\mathcal{A}$. Note that both scales are normalized: the maximum area coverage will have a saturated color, and red hue corresponds to the maximum proximity value. Therefore, they provide only relative (not absolute) information of the area coverage and the intensity (proximity) values. This can be changed fixing absolute ranges for the hue-saturation bi-gradient scale.

A first conclusion that can be directly drawn from Figure 7.3 is that structures of Q are mainly surrounded (more saturation), among those of $[A_{ij}]_+$ and Q itself, by structures with low values of $\hat{C}$, which are also closer (red hues). Those $\hat{C}$ values correspond to sheet-like structures. A desaturated

region (less area coverage) of also green/cyan hues (farther structures) is located nearer the $(\hat{S}, \hat{C}) \approx (1/2, 1)$ zone. This implies that, secondarily, tube-like structures surround also structures of Q, but they are not so proximal and cover a smaller relative surface area of them. When Figure 7.3 is compared to the right plot of Figure 6.4, since both share the same axes $(\hat{S}, \hat{C})$, it is indirectly concluded that the majority of sheet-like structures surrounding those of Q (as seen in Figure 7.3 by the saturated red-colored regions) are structures of $[A_{ij}]_+$ since they are the ones with a higher density of glyphs in the corresponding regions in the $(\hat{S}, \hat{C})$-plane of Figure 6.4.

7.3.2 Proximity split by groups through cumulative marginal pdfs

In Figure 7.4 we plot cumulative one-dimensional marginal pdfs obtained from $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow B}$ for increasing group g numbers, $\{f_k(p) = \sum_{g=1}^{k} [\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow B}, k = 1, ..., N_g\}$. The contribution of each group g is represented by a different color between black lines. For $k = N_g$, the result is the marginal pdf in terms of p, $f(p) = f_{N_g}(p)$, represented by a thicker black line.

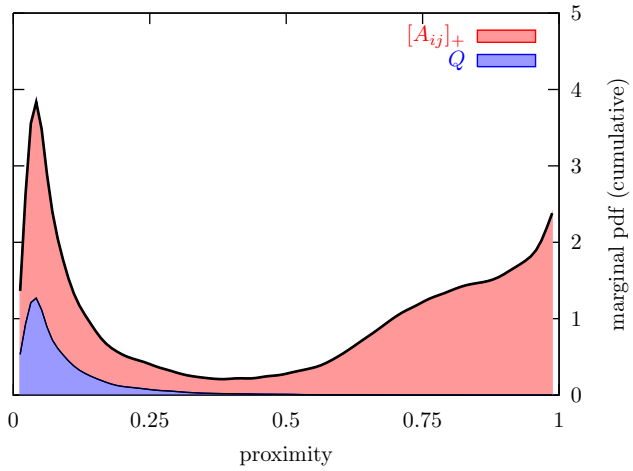


Figure 7.4: Representation of the marginal pdf of $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow B}$ in terms of the proximity p (thick black line), showing the contribution of each group g by the different colored areas between two black lines (added cumulatively)

From Figure 7.4, it is directly concluded that structures of Q are predominantly surrounded by structures of $[A_{ij}]_+$, represented by the red area of the marginal pdf. The blue area corresponds to structures of Q, which cover a much smaller fraction and are farther (lower values of proximity) than those of the majority of $[A_{ij}]_+$ structures. This was indirectly concluded previously, from

the combination of Figures 7.3 and 6.4. The contribution of $[A_{ij}]_+$ (red area) to the marginal pdf in Figure 7.4 shows four regions of interest: first, the region near unitary proximity ($p \approx 1$), that corresponds to structures of $[A_{ij}]_+$ very close to those of Q, likely overlapping/intersecting each other. Examples of this interaction, extracted from the database under study, are shown in Figure 7.5, cases (a), (b), and (c). Second, the region with proximity values between 0.6 and 0.9, where a hump is visible in the marginal pdf, that corresponds to structures of $[A_{ij}]_+$ (or parts of them) still at a close distance but not overlapping or intersecting. Those values of proximity translate into distances between 1/10 and 2/3 times their own characteristic length, defined by μ, which corresponds, for the tube-like structures predominant in Q, to approximately 3/2 times their characteristic radius. See cases (d), (e), and sheets surrounding tubes in case (f) of Figure 7.5 for some examples of this configuration. Third, the region with proximity values between 0.25 and 0.5 with low values of the marginal pdf indicates that surrounding structures are rarely found at those distances. The fourth region corresponds to low values of proximity ($p \approx 0.1-0.2$). See, for example, tubes nearby other tubes in case (f) of Figure 7.5. As a result of the definition of p, a wide range of high values (> 4) of the relative distance between structures concentrates in that region, so it is expectable to find an accumulation of the marginal pdf for low p, whenever there are structures of $\mathcal{A}$ not completely surrounded by close ones of $\mathcal{B}$. A modification of the definition of p or the scale used to represent it can spread the effect of those higher distance over a wider range if required, but here we are more interested in closer structures, so the current p seems suitable for this purpose.

7.3.3 Structures of Q surrounding themselves

A question that arises after this analysis is whether structures of Q might be closer to themselves than what the blue region in Figure 7.4 shows, but structures of $[A_{ij}]_+$ eclipsing them when computing $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right]_{\mathcal{A}\leftarrow\mathcal{B}}$ masquerade the result, making them appear farther than they are. This can be answered by applying the proposed methodology to $\mathcal{A} = \mathcal{B} = \mathcal{X}(Q)$, that is, considering the problem of how structures of Q surround themselves. Figure 7.6 shows the $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right]_{\mathcal{A}\leftarrow\mathcal{B}}$ (left) and $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$ (right) obtained in this case. Note that now there is only one group ($g = 1$).

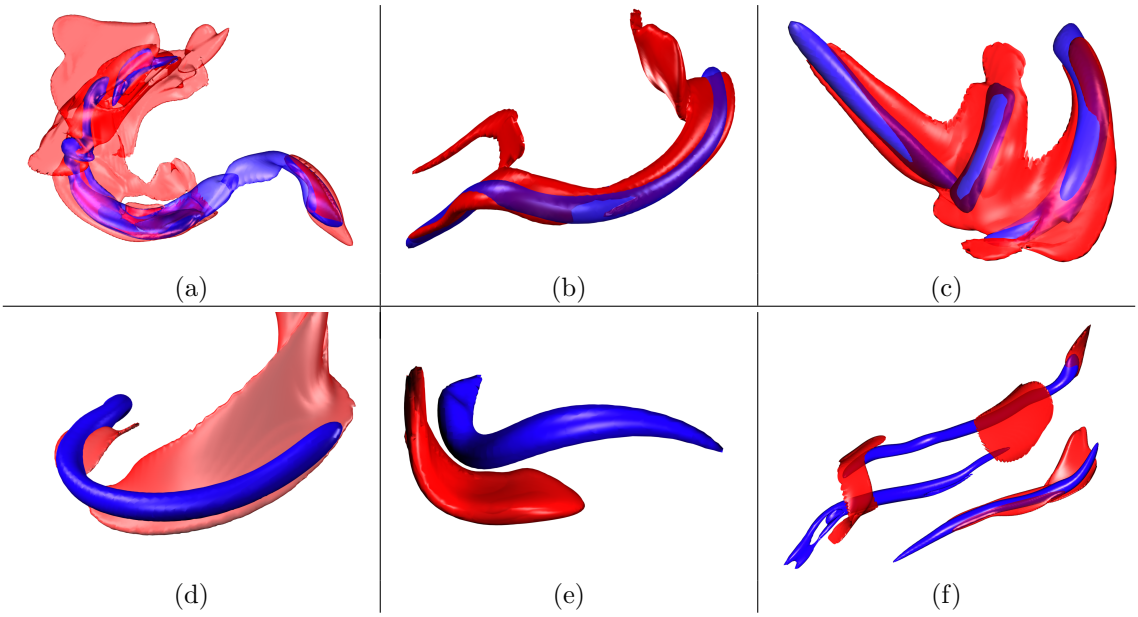


Figure 7.5: Examples of composite structures formed by interaction of tube-like (blue) and sheet-like (red) geometries found in the database. In cases (a) and (b) a tube is mostly embedded by one or more sheets that intersect it and remain at a close distance. Case (c) consists of three tubes and a sheet that follows closely their geometry, intersecting them, and connects them through stretched regions. Cases (d) and (e) show, each, a sheet-like structure wrapping around a tube, without intersecting it but remaining at a close distance and following its curvature. A smaller sheet intersecting the tube is also seen in case (d). Case (f) shows three tubes at moderate distances from each other ($\approx$ 5–10 times their average radius) with a similar orientation and sheets partially surrounding them at close distance or even intersecting them. For clarity, only a subset of all the nearby structures surrounding each tube in every case is shown

$\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right]\Big|_{\mathcal{A}\leftarrow\mathcal{B}}$ shows, as expected, tube-like geometries as the proximal, which is trivial once the sheet-like structures of $[A_{ij}]_+$ have been removed from $\mathcal{B}$. But $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$ also shows a peak at about the same value of proximity that was found for Q structures when $[A_{ij}]_+$ structures were included, which confirms the farther distances among Q structures to themselves. See case (f) in Figure 7.5.

7.4 Application to structures of $\omega_i\omega_i$ and $S_{ij}S_{ij}$

Next we apply this methodology to structures of $\omega_i\omega_i$ and $S_{ij}S_{ij}$ educed in §5.3. For these two fields a multi-scale decomposition was performed. Of all possible combinations of fields and scales, we study two cases, due to their particular relevance. In both cases, we take $\mathcal{A}$ as the set of structures of $\omega_i\omega_i$ at scale number 3 (intermediate scale). They were found in §5.4 to be predominantly tube-

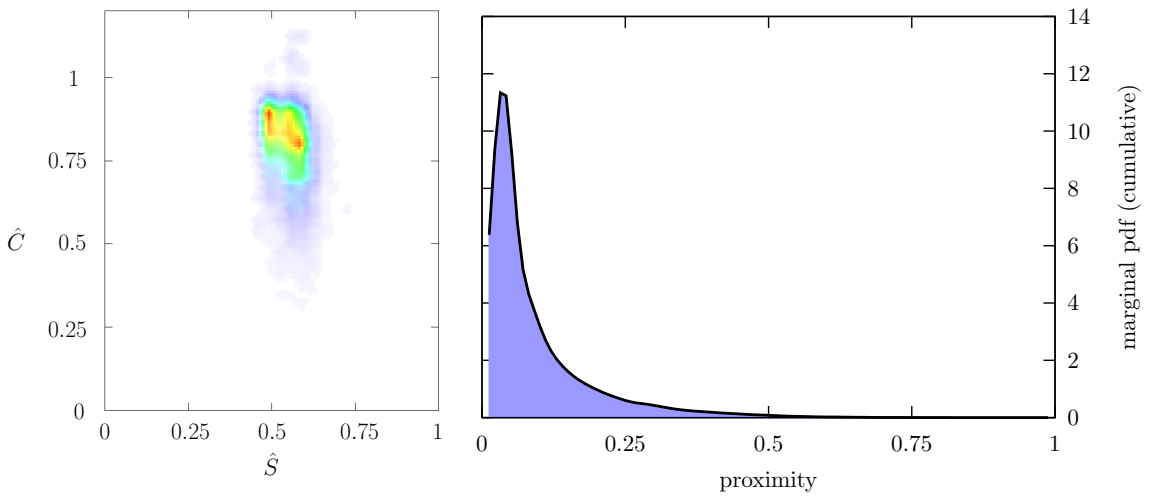


Figure 7.6: Results for structures of Q surrounding themselves: $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right]\Big|_{\mathcal{A}\leftarrow\mathcal{B}}$ (left) $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$ (right) for the case $\mathcal{A}=\mathcal{B}=\mathcal{X}(Q)$, $g=1$

like. We take $\mathcal{B}$ as the set of structures educed for scales numbers $3-6$ for $\omega_i\omega_i$ in the first case (thus, $\mathcal{A}\subset\mathcal{B}$) and for $S_{ij}S_{ij}$ in the second (thus, $\mathcal{A}\cap\mathcal{B}=\emptyset$). In both cases, we split $\mathcal{B}$ into four groups ($N_g=4$), each corresponding to a different scale number of the field under consideration ($\omega_i\omega_i$ or $S_{ij}S_{ij}$, respectively).

Figure 7.7 shows $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right]\Big|_{\mathcal{A}\leftarrow\mathcal{B}}$ (left) and $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$ (right), for the first (top) and second (bottom) cases. From the jpdf+i (left plots) it is observed that structures with small $\hat{C}$ (corresponding to sheet-like geometries) appear to be the closest in both cases (yellow and red spots). But there is a wide range of geometries among the surrounding structures.

In the first case, that is, for the set $\mathcal{B}$ containing structures of $\omega_i\omega_i$, two saturated regions (implying high area-coverage) are present in that plot, corresponding to tube-like (nearer the $(\hat{S},\hat{C})\approx(1/2,1)$ region) and sheet-like (small $\hat{C}$ values) structures. The latter appear closer (red and yellow hues, as opposed to green and cyan) but the former seem to cover a slightly higher percentage of the area of structures of $\mathcal{A}$ (more saturated colors). The region in between those two (intermediate values of $\hat{C}$) contains structures which appear farther and covering a smaller area.

In the second case, $\mathcal{B}$ containing $S_{ij}S_{ij}$ structures, the highest area-coverage corresponds clearly to those structures with low $\hat{C}$ (sheet-like), which are also the closest (red, yellow, and green hues). The spread toward other geometries is also significant, but the area-coverage and proximity decreases

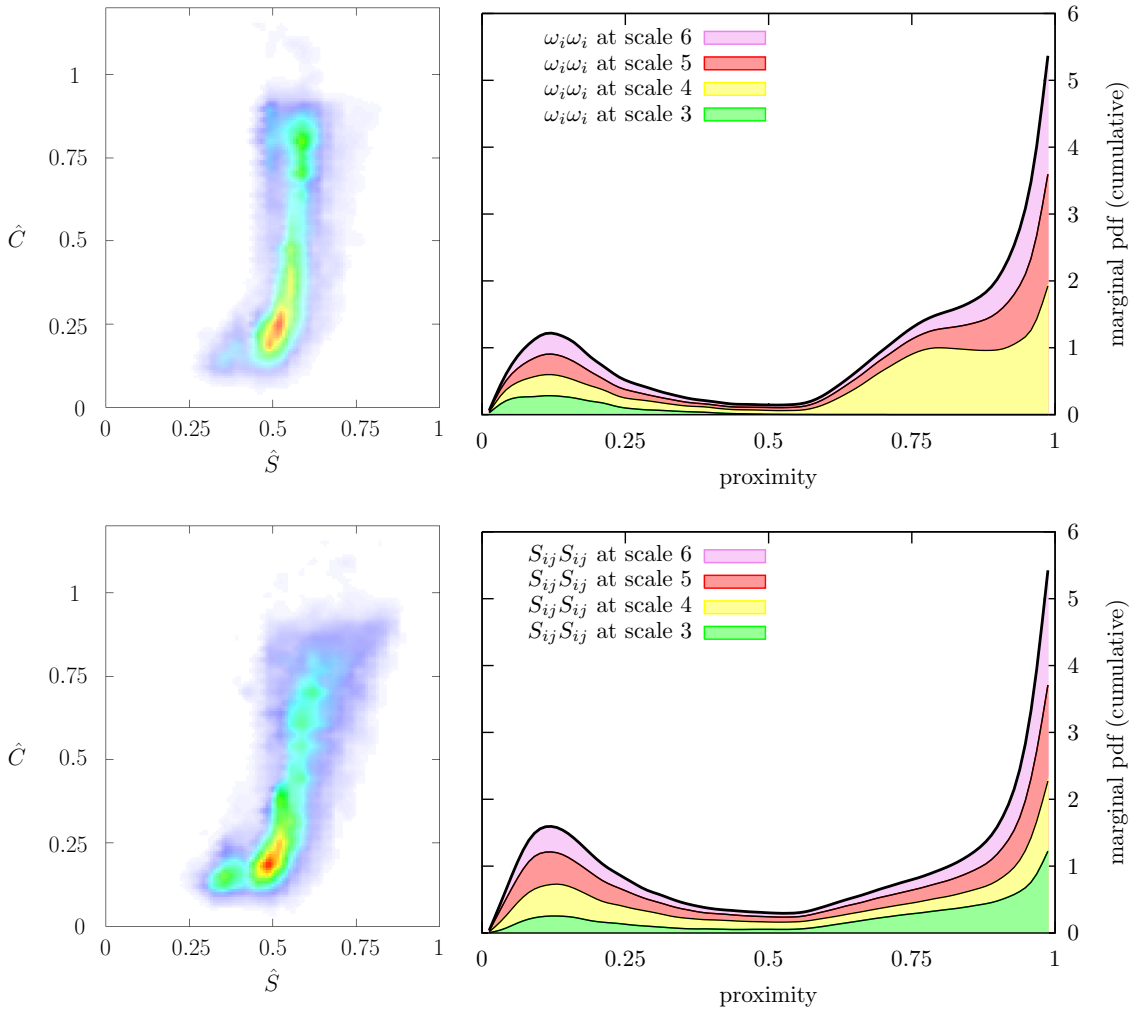


Figure 7.7: Results for structures of $\omega_i\omega_i$ at scale number 3 surrounded by structures of $\omega_i\omega_i$ (top) and by structures of $S_{ij}S_{ij}$ (bottom), at scale numbers $3-6$: $\left[\vec{\mathcal{PI}}(\hat{S},\hat{C}\,;p)\right|_{\mathcal{A}\leftarrow\mathcal{B}}$ (left) $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$ (right)

in those other regions of the $(\hat{S},\hat{C})$-plane (less saturation and cyan and blue hues). Also, the region around tube-like structures found in the first case is now, in the second case, more diffuse and spread toward the blob-like region. This is consistent with the fact that less tubes were found in §5.3 in the geometrical analysis of structures of $S_{ij}S_{ij}$.

The right plots of Figure 7.7 show the cumulative marginal pdf obtained from $[\mathcal{P}(p,g)|_{\mathcal{A}\leftarrow\mathcal{B}}$, split by groups corresponding to each scale number under study (3 to 6). We use the same color code for each scale number as in Figures 5.7– 5.12 of Chapter 5. There are two main differences between both cases of study. In the first case, structures of $\omega_i\omega_i$ at scale number 3 have a small contribution

to the marginal pdf, and they appear far ($p \approx 0.1$), compared to the rest of the scales; structures at scale number 4 are the closest, showing a significant increase for p in the range 0.6-0.8, while structures at scale numbers 5 and 6 appear also close, but with a more gradual increase for $p > 0.6$. In the second case, structures of $S_{ij}S_{ij}$ at each one of the scales under analysis (3-6) have a more balanced contribution to the marginal pdf; scale number 3 shows a slightly higher concentration of closer structures ($p > 0.6$), and less proportion of farther structures ($p \approx 0.1$), when compared to structures at scale numbers 4–6.

7.5 Discussion

From this proximity study, tubes appear closely surrounded by sheets, both when structures of Q and those of an intermediate scale of $\omega_i\omega_i$ are analyzed, in relation to structures of $[A_{ij}]_+$ and the sets of structures of $\omega_i\omega_i$ and $S_{ij}S_{ij}$, respectively. Concerning structures of Q and $[A_{ij}]_+$ surrounding those of Q itself, it is found that a large proportion of structures of $[A_{ij}]_+$ appears much closer, with regions either intersecting or at less than one characteristic diameter of the tube-like structures of Q, and covering a larger proportion of their area than other surrounding structures of Q. These are, on average, farther than five diameters apart from themselves, as is also a second group of $[A_{ij}]_+$ structures, which might be surrounding those other Q structures at a closer distance.

Regarding structures of $\omega_i\omega_i$ at an intermediate scale (scale number 3), which were found to be predominantly tube-like in the previous geometrical study (see Chapter 5), we have considered the surrounding structures at the same and smaller scales (i.e., scale numbers from 3 to 6) of both $\omega_i\omega_i$ and $S_{ij}S_{ij}$, each field independently. In the first case, the set of structures of $\omega_i\omega_i$ at scale number 4 is the predominant group surrounding tubes of $\omega_i\omega_i$ at scale number 3. Structures of smaller scales (i.e., scale numbers 5 and 6) are also found close to those at scale number 3. In comparison, structures at scale number 3 appear farther among themselves. Concerning geometries of proximal structures, we find two predominant groups: sheet-like structures, which appear closer on average, and tube-like structures, certainly farther but with a high proportion of the total area-coverage. Other geometries, intermediate between tubes and sheets, had also a contribution, but to a lesser

degree of both proximity and area coverage.

In the second case, when structures of $S_{ij}S_{ij}$ at scale numbers 3–6 surrounding structures of $\omega_i\omega_i$ at scale number 3 are studied, all scales show similar results of proximity and area coverage. Structures at scale number 3 of $S_{ij}S_{ij}$ have a slightly higher value of proximity, likely owing to the similarity of some structures of both $\omega_i\omega_i$ and $S_{ij}S_{ij}$ at that intermediate scale. Regarding the geometry of proximal structures in this case, they are primarily sheet-like, with significantly higher proximity values and area-coverage. Other geometries can be seen among the surrounding structures but they tend more toward the region of blob-like structures, instead of tube-like structures, as opposed to the case of surrounding structures of $\omega_i\omega_i$ itself. Also, these other geometries are found farther and covering a smaller area fraction, on average.

Physically, these findings are consistent with the phenomenology of tubes being generated by one or multiple proximal sheets discussed in §5.6. Furthermore, as Ruetsch & Maxey (1992) pointed out when studying the evolution of small-scale structures in incompressible homogeneous isotropic turbulence, vortex tubes and vortex sheets should not be considered as separate, independent structures. Instead, they form composite structures, as those shown in Figure 7.5, with dependent geometries.

The results of this proximity analysis also support the structure-based explanation of intermittency, previously explored by Moisy & Jiménez (2004) using box counting methods. They found that intense structures form clusters of inertial-range extent. In our case, the low values of the cumulative marginal pdfs encountered for intermediate values of proximity suggest the existence of empty regions (in the iso-contour context) in physical domain in between composite structures.

Chapter 8

Conclusions and Future Work

8.1 Geometry of individual structures

A methodology for the identification of structures based on their geometry has been introduced first. Our goal has been to develop a methodology that can compensate for the computational bottleneck of DNS computing for turbulent flows, and to provide a solid mathematical framework for non-local characterization of the flow structures based on existing data sets. The main characteristics of this methodology, in comparison with previously existing ones, are its multi-scale and non-local character. The multi-scale nature, implemented by means of the curvelet transform, provides the framework for studying the evolution of the structures associated with the main ranges of scales defined in Fourier space, while keeping the localization in physical space that enables a geometrical study of such structures. We note that the multi-orientation decomposition included in the curvelet transform, not used in this study, can be useful when analyzing other flows in which the directionality of the structures can play a significant role, such as channel flow. The non-local character of the methodology is achieved through the calculation of area-based probability functions of the differential-geometry properties of the surface under consideration. It is also a generic methodology, not intended to educe a particular kind of geometry, but able to manage and classify all possible geometries. There are three main steps involved: extraction, characterization, and classification of structures. Individual structures (considered as closed surfaces disconnected from each other) are studied.

Results of its application first to a virtual world of modeled structures for system validation purposes and then to different fields obtained from DNS turbulence databases have been presented. From its application to the passive scalar fluctuation field advected and diffused in incompressible statistically stationary homogeneous isotropic turbulence in a 513^3 periodic box, the following conclusions can be drawn: first, the multi-scale decomposition resulted in a set of scalar fields (associated with the different ranges of scales extracted) with volumetric probability density functions of decreasing width for smaller scales. In addition, those probability density functions corresponding to scales approximately in the inertial range tend to overlap. Secondly, the study of the structures educed for the different scales shows a transition of their geometry from predominantly the blob-like and tube-like kind in the inertial range of scales toward sheet-like structures in the dissipation range. The dominant structures become more and more stretched for the smaller scales. This transition of geometry is smooth, complicating the automatic classification of structures. There are not clearly distinct groups of structures with a common geometry, but a continuous distribution of them filling the spectrum of present geometries instead. Thus, the application of the clustering algorithms is more challenging. In this case, three groups were educed automatically by applying the clustering technique implemented, and their projection in the visualization space and the identified cluster centers agree with the comments stated above. Nevertheless, clustering results are to be used with care in these conditions in which the points are so continuously distributed in the feature space used for clustering.

This methodology was then applied to the enstrophy and dissipation fields of a second database, obtained from a DNS of incompressible homogeneous isotropic turbulence decaying in time in a periodic box, at the time of maximum enstrophy of the flow. Three different grid resolutions were analyzed, corresponding to 256^3, 512^3, and 1024^3 points, with identical initial conditions and similar $Re_\lambda \approx 77$, resulting in $k_{\max}\bar{\eta}$ of approximately 1, 2, and 4, respectively. This allowed us to compare the geometry of the structures for different resolutions and evaluate whether the traditional DNS grid-resolution criterion $k_{\max}\bar{\eta} \approx 1$ is adequate for such geometrical analysis of the educed structures.

The 1024^3 case showed a continuous transition, for decreasing scale, from blob-like and moder-

ately stretched tube-like structures at large scales to highly stretched sheet-like structures at the smallest scales under study. Intermediate scales of $\omega_i\omega_i$ show a dominance of tube-like structures, which is consistent with the presence of so-called 'worms' in previous studies (Siggia, 1981; Jiménez et al., 1993). Tube-like structures appear also at intermediate scales of $S_{ij}S_{ij}$ but in less proportion than for $\omega_i\omega_i$. The case with smallest grid resolution (256^3 points) did not capture the predominance of highly stretched sheet-like structures educed for the small scales at higher grid resolutions. This suggests the necessity to resolve sub-Kolmogorov scales for a proper geometrical study of the smallest structures of intermittent fields in turbulence, as was previously stated in the literature (see Shumacher & Sreenivasan, 2005; Horiuti & Fujisawa, 2008).

For the 1024^3 case, clustering techniques used during the classification step to obtain distinct groups of geometries among the educed structures resulted in three and two as the optimum number of groups obtained for $\omega_i\omega_i$ and $S_{ij}S_{ij}$, respectively. Blobs, tubes, and sheets can be seen as the predominant structures in the three groups of $\omega_i\omega_i$, while blobs and sheets are predominant in $S_{ij}S_{ij}$; but tubes, present also in this latter field, were included among the two optimum groups. Optimality scores for other number of groups did not differ substantially from the optimal results. This is a consequence of the continuous distribution of geometries, which indicates that the educed groups are not highly differentiated from each other and that the clustering results in this case should again be considered with reserve.

8.2 Assessment of non-local methodology complementing existing local identification criteria

We then applied the same non-local methodology for the study of the geometry of structures to two scalar fields, Q and $[A_{ij}]_+$, used by local criteria of identification of tubes and sheets in turbulence, based on the physical meaning of those quantities, that have been proposed in the fluid mechanics literature. This application confirmed the geometrical character expected for the majority of structures educed from those two fields (which before had been done only visually) by providing the

necessary mathematical and geometrical background as well as means for quantifying the frequency of appearance of each geometry. Clustering techniques in this case provided a much clearer optimum number of two groups of structures, well differentiated. 96% of the structures of Q and $[A_{ij}]_+$ were assigned to separate groups by the clustering algorithm. Q structures were found to be mainly tube-like, while $[A_{ij}]_+$ were recognized as sheet-like. A small amount of structures of both fields present a geometry that does not correspond to the expected shape. For example, some tubes were found among structures of $[A_{ij}]_+$.

8.3 Proximity issues from a geometrical perspective

Finally, we introduced a new methodology for the study of proximity issues among structures corresponding to different fields, from a geometrical perspective. It provides information about the type of geometry found in structures of one group surrounding those of another, indicating the proximity and area coverage, by means of joint probability density functions. The set of surrounding structures can be also split into groups, and quantitative results for each group, concerning the proximity to the other structures, are shown by means of cumulative marginal probability density functions. The representation of the geometrical character of each structure is closely related to the visualization space used in the classification step of the previous study of the geometry of structures in turbulence (as introduced in Chapter 2). We applied this new technique to structures of Q, $[A_{ij}]_+$, $\omega_i\omega_i$ and $S_{ij}S_{ij}$, taken by pairs.

Structures of Q appear closely surrounded, partially overlapped and/or intersected by those of $[A_{ij}]_+$. Comparatively, other structures of Q appear farther from themselves and cover a smaller proportion of their area. A second group of proximal structures of $[A_{ij}]_+$ surrounds those of Q at a farther distance, comparable to the distances where other structures of Q are located, which they might be closely surrounding.

Considering only structures of $\omega_i\omega_i$ those extracted at an intermediate scale (predominantly tube-like) are surrounded primarily by $\omega_i\omega_i$ structures at the immediately smaller scale, and to a lesser degree by structures of even smaller scales. Structures of $\omega_i\omega_i$ at the same intermediate

scale appear significantly farther. Two groups of surrounding geometries are dominant: sheet-like structures are closer; tube-like structures are farther but they cover a large proportion of the area of the structures they surround, thus indicating that the close sheet-like structures are not eclipsing them.

When $\omega_i\omega_i$ at the same intermediate scale are studied in relation to the structures of $S_{ij}S_{ij}$ at that and smaller scales, a more balanced contribution from all scales is observed. Sheet-like geometries are again the closest, and they appear to wrap around the tubes of $\omega_i\omega_i$, eclipsing more effectively other farther geometries.

8.4 Computational remarks

The requirements of our implementation and application of these methodologies to the 256^3 and 512^3 databases do not exceed the computational resources offered by a normal desktop or laptop computer. For the case of 1024^3 grid points, those steps involving Fourier transforms—computation of the fields in physical space from their spectral counterparts, spectral differentiation, and curvelet-based filtering during the multi-scale decomposition—required parallelization and the use of clusters of computers. The rest of the algorithms involved in both methodologies were designed to operate both in parallel and serial environments, independently of the size of the database. For serial operation, splitting and reconnecting algorithms were developed.

8.5 Future work

Both methodologies presented here could benefit from the addition of other geometrical and non-geometrical (e.g., physical) parameters in their analysis. The former could improve the characterization and classification of individual structures, while the latter could be used to relate geometrical properties of those structures with their own physical aspects or those of the surrounding structures. This potential for expansion was a driving criterion during the design and development of both methodologies, translated into the modular character of their implementation.

In the case of the methodology for the study of the geometry of structures, this modularity, consequence of its conceptual division into the three main steps of extraction, characterization, and classification, should facilitate future algorithmic improvements corresponding to each step. For example, the extraction step of the methodology for the study of the geometry of structures currently utilizes iso-surfaces of the (filtered component) scalar fields. Iso-surfaces are a natural first choice to educe structures from a three-dimensional scalar field, but add the dependence on a particular (set of) iso-contour value(s). In our implementation we use the mean of each field plus twice its standard deviation; for higher-contour values the educed structures showed similar geometries, implying a low sensitivity with the contour values, within a certain range. Note that, for extremely high contour values, the number of educed structures will be significantly reduced (only iso-surfaces in the vicinity of a few absolute maxima of the scalar field will be captured) and their geometries could change. Different techniques for the determination of optimum global contour values, such as percolation theory and Morse (critical points) discrete theory were explored in the context of this research. Additional techniques, such as region-based optimum contour values could be applied. Concerning the characterization stage, refined and faster future algorithms for the computation of curvatures of discretized surfaces could be easily implemented. Alternatives to the currently implemented K-means clustering algorithm, part of the classification step, such as *fuzzy c-means* clustering or *density-based* clustering, more oriented toward educing intermingled clusters without clear boundaries can be considered. Also, the addition of other relevant parameters in the clustering process may be useful to allow more separation in those cases where a continuous distribution of geometries is found.

Application to other flows, in particular those with a strong anisotropy, such as channel flow, would be useful, not only to study the geometries present in structures of those flows, but also to compare with homogeneous isotropic turbulence that has been object of this research, as a first canonical case of study. It would help determine whether common geometries in the small scales of turbulence exist. The exploration of the geometry of structures of those anisotropic flows would benefit from the multi-orientation capability of the curvelet transform, part of our current implemen-

tation for the extraction of structures. This capability has not been used in our present applications, for dealing with isotropic turbulence. We note that the mean scalar gradient imposed to the passive scalar field is responsible for some anisotropy in the passive scalar fluctuation field, in contrast with the velocity field that advects it. But due to the relatively weak anisotropy and the use of the passive scalar fluctuation field as the first simpler case to test the methodology, a multi-orientation decomposition was not applied to it.

A multi-orientation analysis could also be useful to study relative alignment among structures, both for isotropic and anisotropic flows. This multi-orientation decomposition can be applied in the extraction of the structures from the original field and then used in the methodology for the study of proximity issues among structures of different sets.

A natural line of expansion of this work is the study of the evolution in time of the geometry of educed structures, as well as the proximity issues presently investigated, both at the individual and composite level. This can be achieved by including a time tracker of structures as an additional module in both methodologies. This may facilitate the search for potential geometrical 'attractors' in a suitable feature space, as well as the development of models of (composite) structure-dynamics. Applications in the tracking and evolution in time of individual Lagrangian structures and their developmental geometry and interaction may also be useful.

Appendix A

Governing equations for the generation of strain, vorticity, dissipation, and enstrophy

A.1 Generation of strain and dissipation

From the Navier-Stokes equations for incompressible flow ($\nabla \cdot \mathbf{u} = 0$):

$$\frac{\partial u_i}{\partial t} + u_k \frac{\partial u_i}{\partial x_k} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_k \partial x_k}, \tag{A.1}$$

decompose the velocity gradient tensor $\partial u_i/\partial x_j$ in its symmetric (strain-rate tensor, S_{ij}) and anti-symmetric (rotation-rate tensor, Ω_{ij}) parts:

$$\frac{\partial u_i}{\partial x_j} = S_{ij} + \Omega_{ij}, \qquad S_{ij} \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad \Omega_{ij} \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right). \tag{A.2}$$

Take $[\partial(A.1)_i/\partial x_j + \partial(A.1)_j/\partial x_i]/2$ and use

$$\frac{1}{2}\left(\frac{\partial u_k}{\partial x_j}\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\frac{\partial u_j}{\partial x_k}\right) = S_{ik}S_{kj} + \Omega_{ik}\Omega_{kj} \tag{A.3}$$

to obtain:

$$\frac{\mathrm{D}S_{ij}}{\mathrm{D}\,t} = -S_{ik}S_{kj} - \Omega_{ik}\Omega_{kj} - \rho^{-1}\frac{\partial^2 p}{\partial x_i \partial x_j} + \nu \frac{\partial^2 S_{ij}}{\partial x_k \partial x_k}, \tag{A.4}$$

where $\mathrm{D}/\mathrm{D}\,t$ denotes the substantial derivative operator ($\mathrm{D}/\mathrm{D}\,t = \partial/\partial t + u_k \partial/\partial x_k$).

Multiply equation A.4 by S_{ji} and use its symmetry to obtain:

$$\frac{\mathrm{D}}{\mathrm{D}\,t}\left(\frac{1}{2}S_{ij}S_{ij}\right) = -S_{ik}S_{kj}S_{ji} - \Omega_{ik}\Omega_{kj}S_{ji} - \rho^{-1}S_{ji}\frac{\partial^2 p}{\partial x_i \partial x_j} + \nu\, S_{ji}\frac{\partial^2 S_{ij}}{\partial x_k \partial x_k}. \tag{A.5}$$

To find the equivalent relations in terms of the vorticity, $\boldsymbol{\omega} \equiv \nabla \times \mathbf{u}$, instead of the rotation-rate tensor, $\boldsymbol{\Omega}$, use $\Omega_{ij} = -\epsilon_{ijk}\,\omega_k/2$, where ϵ_{ijk} is the Levi-Civita symbol and the contraction epsilon identity, $\epsilon_{lik}\,\epsilon_{jmk} = \delta_{lj}\delta_{im} - \delta_{lm}\delta_{ij}$, to express

$$\Omega_{ik}\Omega_{kj} = (\omega_i\omega_j - \delta_{ij}\,\omega_m\omega_m)/4. \tag{A.6}$$

Thus, equations A.4 and A.7 can be rewritten as:

$$\frac{\mathrm{D}S_{ij}}{\mathrm{D}\,t} = -S_{ik}S_{kj} - \frac{1}{4}\left(\omega_i\omega_j - \delta_{ij}\omega_k\omega_k\right) - \rho^{-1}\frac{\partial^2 p}{\partial x_i \partial x_j} + \nu\,\frac{\partial S_{ij}}{\partial x_k \partial x_k}, \tag{A.7}$$

$$\frac{\mathrm{D}}{\mathrm{D}\,t}\left(\frac{1}{2}\,S_{ij}S_{ij}\right) = -S_{ik}S_{kj}S_{ji} - \frac{1}{4}\,\omega_i S_{ij}\omega_j - \rho^{-1}S_{ji}\frac{\partial^2 p}{\partial x_i \partial x_j} + \nu\, S_{ji}\frac{\partial S_{ij}}{\partial x_k \partial x_k}, \tag{A.8}$$

where the symmetry of S_{ji} and the incompressibility condition ($\nabla\cdot\mathbf{u} = \mathrm{tr}(\nabla\mathbf{u}) = \mathrm{tr}(\mathbf{S}) = \delta_{ij}S_{ji} = 0$) have been considered.

A.2 Generation of vorticity and enstrophy

Apply the curl operator to equation A.1, obtaining:

$$\frac{\partial(\nabla \times \mathbf{u})}{\partial t} + \nabla \times (\mathbf{u}\cdot\nabla\mathbf{u}) = -\nabla \times (\nabla p) + \nu\,\nabla^2(\nabla \times \mathbf{u}). \tag{A.9}$$

From tensorial algebra, the following identities hold:

$$\nabla \times (\nabla \phi) \equiv 0, \qquad \nabla \cdot (\nabla \times \mathbf{u}) \equiv 0, \qquad \nabla u^2/2 \equiv \mathbf{u} \cdot \nabla \mathbf{u} + \mathbf{u} \times \nabla \times \mathbf{u}, \tag{A.10}$$

$$\nabla \times (\boldsymbol{A} \times \boldsymbol{B}) \equiv (\boldsymbol{B} \cdot \nabla + \boldsymbol{B} \times \nabla \times)\boldsymbol{A} + (\boldsymbol{A} \cdot \nabla + \boldsymbol{A} \times \nabla \times)\boldsymbol{B}, \tag{A.11}$$

where ϕ is a scalar field, $\mathbf{u}$ is a vector field with modulus $u \equiv \|\mathbf{u}\|$, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are second-order tensor fields. Use these tensor identities and the incompressibility condition, $\nabla \cdot \mathbf{u} = 0$ to rewrite equation A.9 as:

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \mathbf{u} \cdot \nabla \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \nabla \mathbf{u} + \nu \, \nabla^2 \boldsymbol{\omega}. \tag{A.12}$$

The first term of the right-hand side of equation A.12 is responsible for the vortex-stretching.

Dot product equation A.12 with $\boldsymbol{\omega}$, decompose the velocity gradient tensor as $\nabla \mathbf{u} = \boldsymbol{S} + \boldsymbol{\Omega}$ in the first term of the right-hand side, and use the antisymmetry of $\boldsymbol{\Omega}$ [1] to obtain:

$$\frac{\mathrm{D}}{\mathrm{D}\, t}\left(\frac{1}{2}\, \omega_i \omega_i\right) = \omega_i S_{ik} \omega_k + \nu\, \omega_i \frac{\partial^2 \omega_i}{\partial x_k \partial x_k}. \tag{A.13}$$

To express equation A.13 in terms of $\boldsymbol{\Omega}$, particularize

$$\Omega_{ij} \frac{\partial^2 \Omega_{kj}}{\partial x_p \partial x_p} = \frac{1}{4}\, \epsilon_{ijl}\, \omega_l\, \epsilon_{kjm}\, \frac{\partial^2 \omega_m}{\partial x_p \partial x_p} = \frac{1}{4}\left(\omega_m \frac{\partial^2 \omega_m}{\partial x_p \partial x_p}\, \delta_{ik} - \omega_k \frac{\partial^2 \omega_i}{\partial x_p \partial x_p}\right) \tag{A.14}$$

for $k = i$, obtaining:

$$\Omega_{ij} \frac{\partial^2 \Omega_{ij}}{\partial x_p \partial x_p} = \frac{1}{2}\, \omega_i \frac{\partial^2 \omega_i}{\partial x_p \partial x_p}. \tag{A.15}$$

Then, from equation A.6 and the antisymmetry of $\boldsymbol{\Omega}$ it results $\omega_i \omega_i = 2\, \Omega_{ij} \Omega_{ij}$, $\omega_i \omega_j = 4\, \Omega_{ik} \Omega_{kj} + 2\, \delta_{ij} \Omega_{mn} \Omega_{mn}$, which can be substituted, along with the relation A.15, into equation A.13 to obtain:

$$\frac{\mathrm{D}}{\mathrm{D}\, t}\left(\frac{1}{2}\, \Omega_{ij} \Omega_{ij}\right) = 2\, \Omega_{ik} \Omega_{kj} S_{ji} + \nu\, \Omega_{ij} \frac{\partial^2 \Omega_{ij}}{\partial x_k \partial x_k}, \tag{A.16}$$

[1] If $\boldsymbol{A}$ is an antisymmetric second-order tensor, then $\boldsymbol{a} \cdot \boldsymbol{A} \cdot \boldsymbol{a} = 0$, for any vector field $\boldsymbol{a}$.

where the symmetry of $\boldsymbol{S}$ and the incompressibility relation expressed in terms of $\boldsymbol{S}$, $\delta_{ij}S_{ij} = 0$, have been used to rewrite the first term of the right-hand side.

Appendix B

Interpretation of extracted structures

Here we give an interpretation of the physical meaning of the educed structures at different scales resulting from the extraction step of the present methodology. For simplicity and clarity, a two-dimensional scalar field is used. A 1024×1024 greyscale image (with values in the range [0–255]) (see top left of Figure B.1) obtained from a particular realization of a Julia set has been chosen as the scalar field. This set is of interest since it contains self-similar structures at different scales. The outer region has been faded to white so that all boundaries have the same value.

The extraction procedure described in the main text is applied to the two-dimensional image. The result of the multi-scale decomposition provided by the curvelet transform can be seen on the left images of Figure B.1. The effect in Fourier space is shown by the spectra on Figure B.2(b), while Figure B.2(a) shows the bin pdfs of the original and filtered fields, in physical space. Note that, in this case, the low-pass filter used for the coarsest scale is (in logarithmic scale) wider than the others. It can be thought of as two scales merged into one (the coarsest scale, in this case), and could be done also for other groups of scales.

Each filtered field (image) corresponding to each scale is then iso-contoured at a value equal to its mean plus 3/2 times its standard deviation: see right plots on Figure B.1. The original field has also been iso-contoured (top right) for comparison.

From the way in which the decomposition is done, as observed in the spectra, the structures educed for each filtered scale have a correspondence with the different 'energetic' bands of the

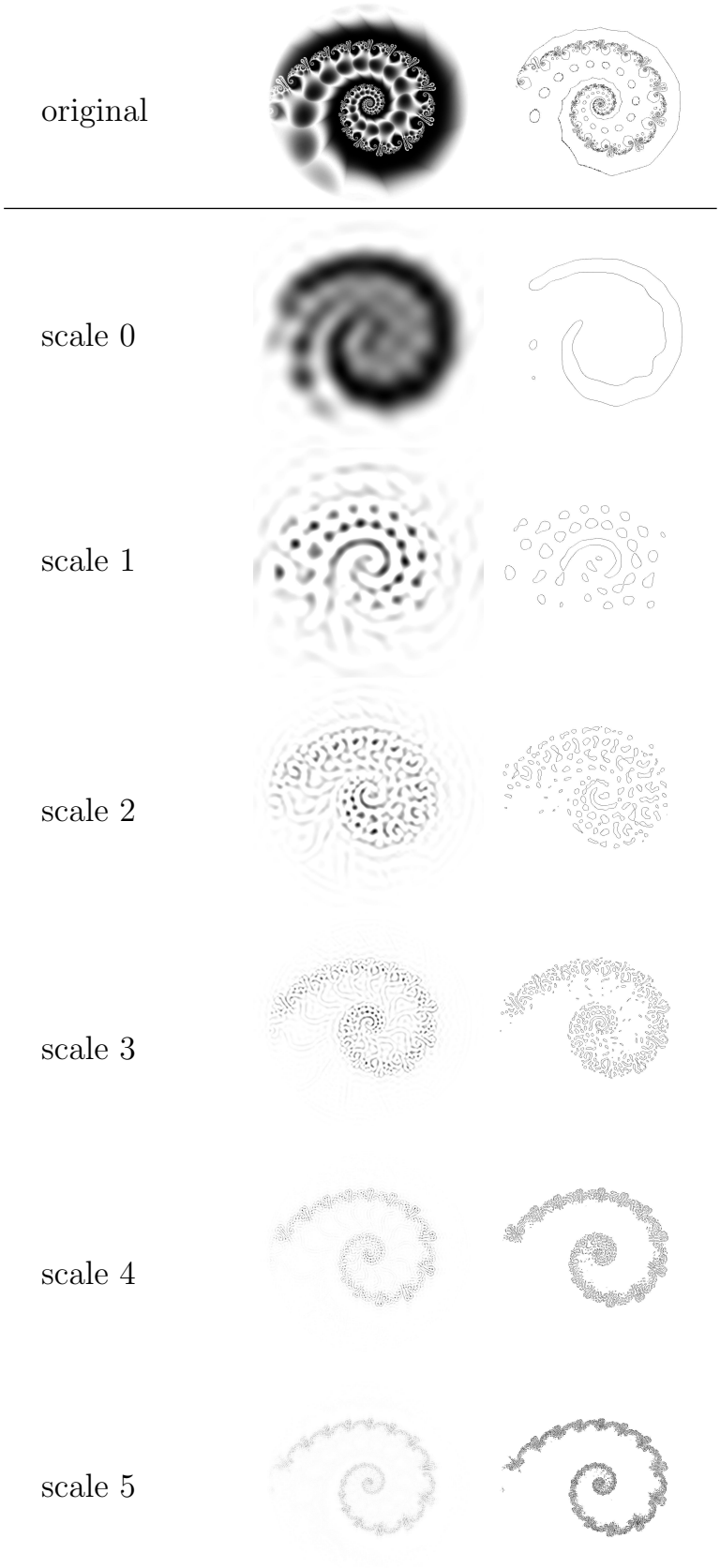


Figure B.1: Fields (left) and corresponding iso-contours (right) for original (top) and filtered scales (below)

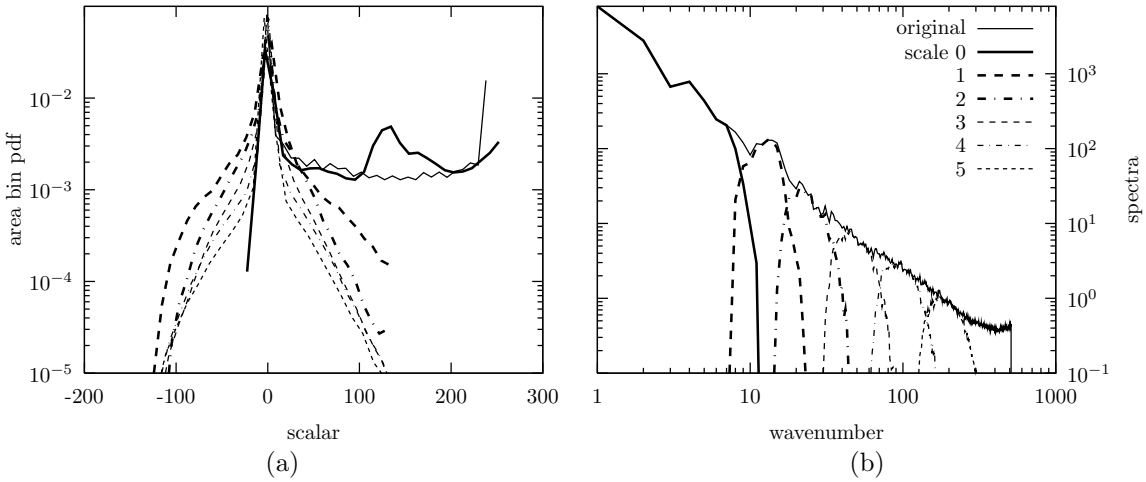


Figure B.2: Pdfs (a) and spectra (b) associated with the original and each one of the filtered scales

field, as defined by frequency corona in Fourier space. But furthermore, they also have a direct correspondence in physical space with the structures of the original field (features of the image). First, we notice that the spatial localization of the features (structures) educed for each scale is retained, with respect to the original image. As expected, they vary in relative sizes (scales), from one filtered scale to the next. Some features of the original image that span across different scales are split as a result of the decomposition. See, for example, the dark continuous arm of the spiral: scale 0 captures its largest portion, but the remainder can be seen also in the rest of the scales. The geometry of each part resembles that of the structure from which it was derived. Shape is preserved and thus a geometrical analysis of the educed structures is meaningful in this context. The iso-contour obtained from the original field (top right of Figure B.1) contains a large individual structure, rich in features, and a few simpler structures, but is missing many other features of the original image. In contrast, contours of the filtered fields tend to contain many more (simpler) structures that capture the essential features of the original field at that scale. The fact that the spatial localization is kept can be used for the study of relative positioning, clustering, and other organizational aspects of the sets of structures.

The extension of this reasoning to three dimensions is immediate. The complexity of the structures that can be found increases. For example, structures that appear as circular in two dimensions could become either blob- or tube-like, while elongated structures in two dimensions could become

either tube- or sheet-like structures. We note that an alternative to the multi-scale decomposition of the scalar field applied here is to perform a multi-resolution analysis applied to the iso-contours extracted from the original database. Since there is a loss of information by iso-contouring, we choose to perform the multi-scale decomposition first over the entire field and then iso-contour each one of the filtered scales.

As an analogy, consider the decomposition of a tree into its trunk, branches, leaves, etc. The outer surface of the tree, containing all those elements, would correspond to the iso-surface of the original field. It is generally too rich and complex to study as a whole. By applying a multi-scale decomposition before iso-contouring, we can separate the tree into its individual components, ranged by the scale. Then, iso-contouring extracts structures at those different levels, whose properties can be studied individually. This is the philosophy applied in our methodology for the study of structures in turbulence. In the same manner that the geometry of the elements of a tree has a relation to their physical functionality, perhaps that is also the case for those structures present in turbulent flows. A multi-scale decomposition followed by surface identification (by iso-contouring based on global contour values or other means) seems an appropriate framework for this study. Its current form can be considered a starting point, but there is much room for refinement: for example, use of additional multi-resolution capabilities (as outlined in the body of the thesis), such as multi-orientation decomposition, and selection of locally adapted contour levels for optimal feature extraction are two possible paths for improvement.

Appendix C

Differential geometry background

Let M be a *regular surface*[1] parametrized by:

$$\boldsymbol{x}(\boldsymbol{u}) : \{x(u,v),\, y(u,v),\, z(u,v)\}. \tag{C.1}$$

Its *tangent plane* at any point P, T_PM, is defined by the tangent vectors $\{\boldsymbol{x}_u,\, \boldsymbol{x}_v\}$ contained on it, or, alternatively, by the normal unit vector $\mathbf{N}$ orthogonal to it:

$$\boldsymbol{x}_u = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right), \qquad \boldsymbol{x}_v = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right), \tag{C.2}$$

$$\boldsymbol{N} = \frac{\mathbf{x}_u \wedge \mathbf{x}_v}{||\mathbf{x}_u \wedge \mathbf{x}_v||}. \tag{C.3}$$

The *first fundamental form* of M at a point P is the inner product restricted to tangent vectors:

$$I(\boldsymbol{m}_P, \boldsymbol{n}_P) = \boldsymbol{m}_P \cdot \boldsymbol{n}_P, \tag{C.4}$$

where $\boldsymbol{m}_P$, $\boldsymbol{n}_P \in T_PM$ (tangent plane of M at P). The first fundamental form is independent of the surface representation, and therefore invariant under parameter transformations. It satisfies:

$$I(a\,\boldsymbol{x}_u + b\,\boldsymbol{x}_v, a\,\boldsymbol{x}_u + b\,\boldsymbol{x}_v) = E\,a^2 + 2\,F\,a\,b + G\,b^2, \tag{C.5}$$

[1] $M \subset \mathbb{R}^n$ is a *regular surface* if for each point $P \in M$ there exists a neighborhood of P, $V \in \mathbb{R}^n$, and a map $x : U \to M$ of an open set $U \subset \mathbb{R}^2$ onto $V \cap M$ such that: (i) x is differentiable; (ii) $x : U \to V \cap M$ is a homeomorphism; (iii) each map $x : U \to M$ is a regular patch, that is, its Jacobian has rank 2 for all $(u, v) \in U$.

where E, F, G are the first fundamental coefficients:

$$E = \boldsymbol{x}_u \cdot \boldsymbol{x}_u = ||\boldsymbol{x}_u||^2, \tag{C.6}$$

$$F = \boldsymbol{x}_u \cdot \boldsymbol{x}_v, \tag{C.7}$$

$$G = \boldsymbol{x}_v \cdot \boldsymbol{x}_v = ||\boldsymbol{x}_v||^2. \tag{C.8}$$

These coefficients are not invariant under parameter transformations. Since $||\boldsymbol{x}_v \cdot \boldsymbol{x}_v|| < ||\boldsymbol{x}_u||||\boldsymbol{x}_v||$ (recall $\boldsymbol{x}_u \nparallel \boldsymbol{x}_v$ in a regular surface for the tangent plane to be defined):

$$E\,G - F^2 = ||\boldsymbol{x}_u||^2||\boldsymbol{x}_v||^2 - (\boldsymbol{x}_v \cdot \boldsymbol{x}_v)^2 > 0, \tag{C.9}$$

and therefore, the first fundamental form is a positive definite quadratic form on the tangent plane of M at P (T_PM):

$$I(\boldsymbol{m}, \boldsymbol{n}) = \begin{pmatrix} m_u & m_v \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} n_u \\ n_v \end{pmatrix}, \qquad \left\{ \begin{array}{c} m = m_u\,\boldsymbol{x}_u + m_v\,\boldsymbol{x}_v \\ n = n_u\,\boldsymbol{x}_u + n_v\,\boldsymbol{x}_v \end{array} \right\} \in T_P. \tag{C.10}$$

A property of the surface M that depends only on the first fundamental form of M is called an *intrinsic property* [2].

The *arc length*, $s(t)$, of a curve C on M (given by its parametrization $\vec{\alpha}(t) = (u(t), v(t)) = u(t)\,\boldsymbol{x}_u + v(t)\,\boldsymbol{x}_v$, $t \in [a, b]$), is:

$$s(t) = \int_a^t ||\vec{\alpha}'(r)||\,\mathrm{d}r, \qquad \vec{\alpha}'(r) \equiv \frac{\mathrm{d}\vec{\alpha}(r)}{\mathrm{d}r}. \tag{C.11}$$

[2] An *intrinsic property* of a surface is independent of the space in which the surface may be considered. Thus, a hypothetical "inhabitant" of the surface can measure it without knowing anything about the space in which the surface is embedded. On the other hand an *extrinsic property* of a surface depends on its embedding space, and therefore cannot be measured by "inhabitants" of the surface. In a more formal definition a property is called *intrinsic* if it is preserved by *local isometries*, and *extrinsic* otherwise. An *isometry* (or *congruence transformation*) is a bijective distance preserving map between two metric spaces.

Thus:

$$\begin{aligned}\left(\frac{\mathrm{d}s(t)}{\mathrm{d}t}\right)^2 &= ||\vec{\alpha}'(t)||^2 = \vec{\alpha}'(t)\cdot\vec{\alpha}'(t)\\ &= [u'(t)\,\boldsymbol{x}_u + v'(t)\,\boldsymbol{x}_v]\cdot[u'(t)\,\boldsymbol{x}_u + v'(t)\,\boldsymbol{x}_v]\\ &= E\,u'(t)^2 + 2\,F\,u'(t)\,v'(t) + G\,v'(t)^2\end{aligned}$$

and:

$$(\mathrm{d}s)^2 = E\,(\mathrm{d}u)^2 + 2\,F\,\mathrm{d}u\,\mathrm{d}v + G\,(\mathrm{d}v)^2 = I(\vec{\alpha}',\vec{\alpha}'). \tag{C.12}$$

Therefore, the arc length is an intrinsic property, since it depends only on the first fundamental form. The *area element*, $\mathrm{d}A$, of M at a point $P \in M$ is defined, in terms of its parametrization, as:

$$\mathrm{d}A = \sqrt{E\,G - F^2}\,\mathrm{d}u \wedge \mathrm{d}v, \tag{C.13}$$

where $\mathrm{d}u \wedge \mathrm{d}v$ is the wedge product.

The *second fundamental form* of a (three-dimensional regular) surface M at a point P is the symmetric bilinear[3] form on the tangent plane at P (T_PM) given by:

$$II(\boldsymbol{m}_P, \boldsymbol{n}_P) = S(\boldsymbol{m}_P)\cdot\boldsymbol{n}_P = \boldsymbol{m}_P\cdot S(\boldsymbol{n}_P), \tag{C.14}$$

where S is the *shape operator* (or *second fundamental tensor* or *Weingarten map*), which is defined, when operating on a vector $\boldsymbol{m}$, as the negative covariant derivative (along the direction of $\boldsymbol{m}$), D_m, of the unit normal vector field $\boldsymbol{N}$ of the surface M:

$$S(\boldsymbol{m}) = -D_{\boldsymbol{m}}\boldsymbol{N}. \tag{C.15}$$

[3]Equivalently quadratic in this context

The second fundamental form satisfies:

$$II(a\,\boldsymbol{x}_u + b\,\boldsymbol{x}_v, a\,\boldsymbol{x}_u + b\,\boldsymbol{x}_v) = e\,a^2 + 2\,f\,a\,b + g\,b^2, \tag{C.16}$$

where e, f, g are the second fundamental coefficients:

$$\begin{aligned} e &= -\boldsymbol{N}_u \cdot \boldsymbol{x}_u = \boldsymbol{N} \cdot \boldsymbol{x}_{uu} = \frac{\det(\boldsymbol{x}_{uu}, \boldsymbol{x}_u, \boldsymbol{x}_v)}{\sqrt{E\,G - F^2}}, && \text{(C.17)} \\ f &= -\boldsymbol{N}_v \cdot \boldsymbol{x}_u = \boldsymbol{N} \cdot \boldsymbol{x}_{uv} = \frac{\det(\boldsymbol{x}_{uv}, \boldsymbol{x}_u, \boldsymbol{x}_v)}{\sqrt{E\,G - F^2}} && \text{(C.18)} \\ &= -\boldsymbol{N}_u \cdot \boldsymbol{x}_v = \boldsymbol{N} \cdot \boldsymbol{x}_{vu} = \frac{\det(\boldsymbol{x}_{vu}, \boldsymbol{x}_u, \boldsymbol{x}_v)}{\sqrt{E\,G - F^2}}, && \text{(C.19)} \\ g &= -\boldsymbol{N}_v \cdot \boldsymbol{x}_v = \boldsymbol{N} \cdot \boldsymbol{x}_{vv} = \frac{\det(\boldsymbol{x}_{vv}, \boldsymbol{x}_u, \boldsymbol{x}_v)}{\sqrt{E\,G - F^2}}, && \text{(C.20)} \end{aligned}$$

being:

$$\boldsymbol{x}_{\alpha\beta} \equiv \left(\frac{\partial^2 x}{\partial\alpha\partial\beta}, \frac{\partial^2 y}{\partial\alpha\partial\beta}, \frac{\partial^2 z}{\partial\alpha\partial\beta} \right). \tag{C.21}$$

The last equality in equations C.17–C.20 for the coefficients comes from rewriting the normal vector as $\boldsymbol{N} = (\boldsymbol{x}_u \wedge \boldsymbol{x}_v)/||\boldsymbol{x}_u \wedge \boldsymbol{x}_v|| = (\boldsymbol{x}_u \wedge \boldsymbol{x}_v)/\sqrt{E\,G - F^2}$ where the result $||\boldsymbol{x}_u \wedge \boldsymbol{x}_v|| = \sqrt{E\,G - F^2}$ is a consequence of the Lagrange identity $((\boldsymbol{a} \cdot \boldsymbol{b})^2 + ||\boldsymbol{a} \wedge \boldsymbol{b}||^2 = ||\boldsymbol{a}||^2||\boldsymbol{b}||^2)$ and recalling that $(\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$. Unlike the first fundamental form, the second fundamental form is not necessarily positive or definite.

The *normal curvature* of a regular surface M in the direction of a unit tangent vector $\boldsymbol{t}_P$ at a point $P \in M$ is formally defined as:

$$\kappa_N(\boldsymbol{t}_P) = S(\boldsymbol{t}_P) \cdot \boldsymbol{t}_P, \tag{C.22}$$

S being the shape operator. From the previous definition of the second fundamental form, we can express the normal curvature κ_N as:

$$\kappa_N(\boldsymbol{t}_P) = II(\boldsymbol{t}_P, \boldsymbol{t}_P), \tag{C.23}$$

and, for a generic non-unitary tangent vector, $\boldsymbol{t}_P^{'}$:

$$\kappa_N(\boldsymbol{t}_P^{'}) = \frac{S(\boldsymbol{t}_P^{'}) \cdot \boldsymbol{t}_P^{'}}{\boldsymbol{t}_P^{'} \cdot \boldsymbol{t}_P^{'}} = \frac{II(\boldsymbol{t}_P^{'}, \boldsymbol{t}_P^{'})}{I(\boldsymbol{t}_P^{'}, \boldsymbol{t}_P^{'})}. \tag{C.24}$$

The maximum (κ_1) and minimum (κ_2) values of the normal curvature at a point $P \in M$ are called *principal curvatures*. The directions defined by the tangent vectors associated with those principal curvatures are called *principal directions* and are orthogonal. Formally, the principal curvatures at a point P are defined as the eigenvalues (κ_1, κ_2) of the shape operator $S(P)$, and the principal directions correspond to the associated (orthogonal) eigenvectors ($\boldsymbol{e}_1$,$\boldsymbol{e}_2$). The normal curvature at P along $\boldsymbol{t}_P$ is then given by *Euler's formula*:

$$\kappa_N = \kappa_1 \cos^2(\theta) + \kappa_2 \sin^2(\theta), \tag{C.25}$$

where θ is the angle between $\boldsymbol{e}_1$ and $\boldsymbol{t}_P$. The *Gaussian curvature* of a regular surface M at a point $P \in M$ is formally defined as the determinant of the shape operator S at that point:

$$K(P) = \det(S(P)). \tag{C.26}$$

Gauss' *Theorema Egregium* proves that the Gaussian curvature of a regular surface M is invariant under local isometry. In other words, it is an intrinsic property of the surface, and therefore it only depends on its first fundamental form (at every point $P \in M$). This is a remarkable result since the formal definition of the Gaussian curvature involves the second fundamental form directly (and, therefore, the embedding of the surface M). In terms of the first fundamental form only, the Gaussian curvature is written as:

$$K = \frac{1}{\sqrt{E\,G - F^2}} \left[\frac{\partial}{\partial v} \left(\frac{\sqrt{E\,G - F^2}}{E}\, \Gamma_{11}^2 \right) - \frac{\partial}{\partial u} \left(\frac{\sqrt{E\,G - F^2}}{E}\, \Gamma_{12}^2 \right) \right], \tag{C.27}$$

where Γ_{ij}^k are the Christoffel symbols of the second kind, which can be expressed in terms of the

first fundamental coefficients as:

$$\Gamma_{11}^2 = \frac{2\,E\,F_u - E\,E_v - F\,E_u}{2\,(E\,G - F^2)}, \qquad \Gamma_{12}^2 = \frac{E\,G_u - F\,E_v}{2\,(E\,G - F^2)}. \tag{C.28}$$

Gaussian curvature can also be expressed in terms of the first and second fundamental coefficients in a more compact way:

$$K = \frac{e\,g - f^2}{E\,G - F^2}. \tag{C.29}$$

In terms of the principal curvatures, the Gaussian curvature is expressed as:

$$K = \kappa_1 \kappa_2. \tag{C.30}$$

Points with positive/negative Gaussian curvature are called *elliptic/hyperbolic*. A point is *parabolic* if the Gaussian curvature is zero but not the shape operator. At *planar* points both the Gaussian curvature and the shape operator are zero. In a *synclastic/anticlastic* surface, all its points are elliptic/hyperbolic.

The *mean curvature* of a regular surface M at a point $P \in M$ is formally defined as the trace of the shape operator S at that point:

$$H(P) = \mathrm{tr}(S(P)). \tag{C.31}$$

In terms of the first and second fundamental coefficients, the mean curvature is:

$$H = \frac{e\,G - 2\,f\,F + g\,E}{2\,(E\,G - F^2)}. \tag{C.32}$$

Unlike Gaussian curvature, which is intrinsic, the mean curvature is an extrinsic property of the surface, that is, it depends on the embedding [4]. The mean curvature coincides with the mean of the principal curvatures:

$$H = \frac{\kappa_1 + \kappa_2}{2}. \tag{C.33}$$

[4]For instance, a cylinder and a plane are locally isometric but the mean curvature of a plane is zero while that of a cylinder is nonzero. Their Gaussian curvature is zero.

Combining equations C.30 and C.33, the principal curvatures can be obtained from the Gaussian and mean curvatures:

$$\begin{cases} \kappa_1 = H + \sqrt{H^2 - K} \\ \kappa_2 = H - \sqrt{H^2 - K} \end{cases} . \tag{C.34}$$

Gaussian and mean curvature satisfy:

$$H^2 - K = \left(\frac{\kappa_1 - \kappa_2}{2} \right)^2 \geq 0. \tag{C.35}$$

Points where $H^2 = K$ (that is, $\kappa_1 = \kappa_2$ and, therefore, the normal curvature is the same in any direction) are called *umbilical points*. A surface is defined as (locally) *minimal* if its mean curvature (locally) vanishes ($H = 0$, that is, $\kappa_1 = -\kappa_2$).

The *shape index*, Υ, and *curvedness*, Λ, of a regular surface M at a point P are defined (Koenderink & van Doorn, 1992) by:

$$\Upsilon = -\frac{2}{\pi} \arctan \left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \right), \tag{C.36}$$

$$\Lambda = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}. \tag{C.37}$$

In terms of the Gaussian and mean curvatures, the shape index and curvedness can be expressed as:

$$\Upsilon = -\frac{2}{\pi} \arctan(\frac{H}{\sqrt{H^2 - K}}), \tag{C.38}$$

$$\Lambda = \sqrt{2H^2 - K}. \tag{C.39}$$

Shape index is dimensionless, while curvedness has the dimensions of a reciprocal length. The planar patch, for which $\kappa_1 = \kappa_2 = 0$, has null curvedness and an indeterminate shape index. All other regular patches of a regular surface M map on the domain $(\Upsilon, \Lambda) \in [-1, +1] \times \mathbb{R}^+$. $\{\rho, \phi\} \equiv \{\sqrt{2}\Lambda, -\pi\Upsilon/2\}$ are polar coordinates in the semi-plane[5] of axes $\{\kappa_1 - \kappa_2\}^+$ and $\{\kappa_1 + \kappa_2\}$ (see Figure C.1). Some other properties of the shape index are summarized below (extracted

[5]Note that, by definition, $\kappa_1 - \kappa_2 > 0$.

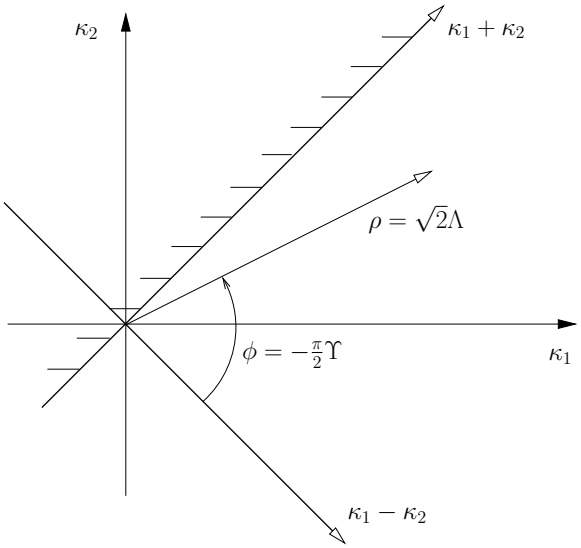


Figure C.1: Transformation from (κ_1, κ_2) to (Υ, Λ)

from Koenderink & van Doorn (1992)) (see Figure C.2 for a graphical explanation):

- Points where $|\Upsilon| = 1$ are umbilical points and represent locally spherical shapes ("cup" ($\Upsilon = -1$) or "cap" ($\Upsilon = +1$))[6]. Points where $0.5 < |\Upsilon| < 1$ are elliptic points and represent locally ellipsoidal shapes, tending toward the spherical shape when $|\Upsilon| \to 1$ and towards the cylindrical shape when $|\Upsilon| \to 0.5$. Points where $|\Upsilon| = 0.5$ are parabolic points and represent cylindrical shapes ("rut" ($\Upsilon = -0.5$), and "ridge" ($\Upsilon = +0.5$)). Points where $0 < |\Upsilon| < 0.5$ are hyperbolic points and represent locally hyperbolic shapes, tending toward the cylindrical shapes when $|\Upsilon| \to 0.5$ and towards the symmetrical saddle when $|\Upsilon| \to 0$.

- The range $\Upsilon \in (-1, -0.5)$ represents the concavities (concave "ruts" or "trough" shapes). The range $\Upsilon \in (-0.5, +0.5)$ represents the saddle-like shapes ("saddle-ruts" ($\Upsilon \in (-0.5, 0)$) and "saddle-ridges" ($\Upsilon \in (0, +0.5)$)), with the symmetrical saddle at $\Upsilon = 0$. The range $\Upsilon \in (+0.5, +1)$ represents the convexities (convex "ridges", or "dome-shapes").

- Generically, umbilicals ($|\Upsilon| = 1$) occur only at isolated points on the surface. Parabolic points ($|\Upsilon| = 0.5$) occur on curves of two distinct types ($\Upsilon = \pm 0.5$), which are smooth, closed

[6]The following convention has been chosen: a regular surface M is locally concave/convex at a given point $P \in M$ if the point P is a local minimum/maximum in the reference system with vertical axis pointing towards the outward normal at P.

loops on closed regular surfaces, and such that never intersect (although they can be nested or juxtaposed). Symmetrical saddles ($\Upsilon = 0$) also occur on curves. Ellipsoid patches with different sign(Υ) ("domes" and "troughs") are never adjacent, being necessarily separated by saddle-like patches.

- Two shapes with opposite shape indices represent complementary pairs (matching each other as "mold" and "stamp", when appropriately scaled).

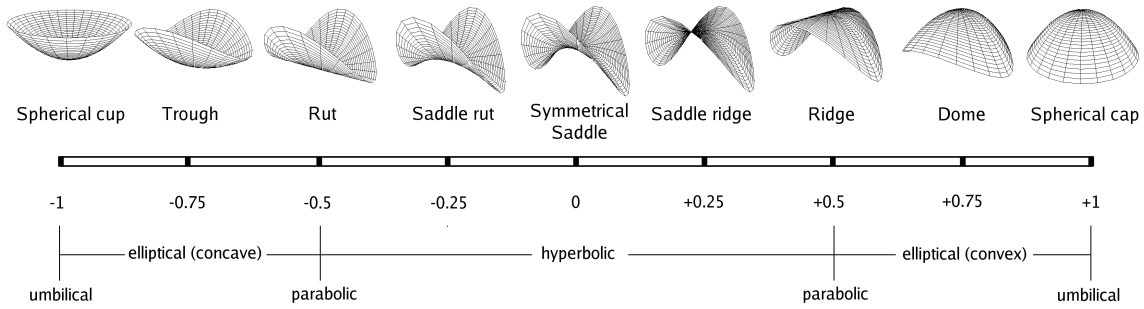


Figure C.2: Range of shape index (Υ), with its most representative associated local shapes (figure based on Koenderink & van Doorn (1992))

Appendix D

Density functions on manifolds

Consider a function $\xi_M(P) : P \in M \mapsto I \subseteq \mathbb{R}$, that defines a local property ξ of an m-dimensional manifold M embedded in a n-dimensional space $\mathbb{R}^n$, $m < n$. At every point $P \in M$. Define $M_\xi^=$ as the set of points on M where $\xi_M(P)$ is equal to a particular value ξ, $M_\xi^= = \{P \in M \mid \xi_M(P) = \xi\}$, and $M_\xi^\leq$ as the set of points on M where $\xi_M(P)$ is less or equal to a particular value ξ, $M_\xi^\leq = \{P \in M \mid \xi_M(P) \leq \xi\}$. Consider the *measure spaces*[1] $(\mathbb{R}^n, \mathcal{F}(\mathbb{R}^n), \mu_1)$ and $(\mathbb{R}^n, \mathcal{F}(\mathbb{R}^n), \mu_2)$, where $\mathcal{F}(\mathbb{R}^n)$ is a σ-algebra of $\mathbb{R}^n$, and μ_1, μ_2 are two particular measures defined on $(\mathbb{R}^n, \mathcal{F}(\mathbb{R}^n))$.

Define the function $\Psi(\xi) : \xi \in I \mapsto \bar{\mathbb{R}}^+$ such that for every value $\xi \in I$ it returns the μ_1-measure of the set $M_\xi^= \subset M$:

$$\Psi(\xi) \equiv \mu_1\left(M_\xi^=\right) \equiv \int_{M_\xi^=} \mathrm{d}\mu_1 = \int_M \mathbf{1}_{[M_\xi^=]} \mathrm{d}\mu_1, \tag{D.1}$$

where the function $\mathbf{1}_{[M_\xi^=]}$ is the *characteristic function*[2] on $M_\xi^=$ and the integrals are defined in the generalized Lebesgue[3] sense.

[1] A *measure space* $(E, \mathcal{F}(E), \mu)$ is a measurable space, $(E, \mathcal{F}(E))$, with a non-negative measure, μ. A *measurable space*, $(E, \mathcal{F}(E))$, is a set E with a σ-algebra, $\mathcal{F}(E)$, on it. A σ-algebra $\mathcal{F}$ on a given set E is a nonempty collection of subsets of E such that: 1) $\emptyset \in \mathcal{F}(E)$; 2) if $A \in \mathcal{F}$ then $\bar{A} \in \mathcal{F}$, where $\bar{A}$ is the complement of A; 3) if A_n is a sequence of elements of $\mathcal{F}$, then $\bigcup A_n \in \mathcal{F}$. As a consequence: $E \in \mathcal{F}$. A *measure* μ, defined on a measurable space $(E, \mathcal{F}(E))$, is a function $\mu : \mathcal{F}(E) \mapsto \bar{\mathbb{R}}$ (where $\bar{\mathbb{R}}$ denotes the extended real numbers ($\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$)) such that: 1) $\mu(A) \geq 0$ for $A \in \mathcal{F}(E)$ (equality iff $A = \emptyset$), 2) $\mu(\cup_{n=0}^\infty A_n) = \sum_{n=0}^\infty \mu(A_n)$ for any sequence of disjoint sets $A_n \in \mathcal{F}(E)$ (countable additivity). If $\mu(E) = 1$ then μ is called a *probability measure* and the measure space $(E, \mathcal{F}(E), \mu)$ is called a *probability space*.

[2] The *characteristic function* or *indicator function*, of a subset $A \in E$ is a function $\mathbf{1}_A : E \mapsto \{0, 1\}$ defined as $\mathbf{1}_A = \{1, \text{if } P \in A; 0, \text{if } P \notin A\}$.

[3] The *Lebesgue integral* of a measurable function $f : E \mapsto \bar{\mathbb{R}}$ on a measure space $(E, \mathcal{F}(E), \mu)$, is defined through the following steps: 1) for the characteristic function, $\mathbf{1}_A$, $\int_E \mathbf{1}_A \, \mathrm{d}\mu \equiv \mu(A)$; 2) for a simple function (i.e., $s = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$, $c_i \in \mathbb{R}$, for some finite collection $A_i \in \mathcal{F}(E)$), then: $\int_E f \, \mathrm{d}\mu \equiv \sum_{i=1}^n c_i \int_E \mathbf{1}_{A_i} \, \mathrm{d}\mu = \sum_{i=1}^n c_i \mu(A_i)$; 3) For a non-negative measurable function f (possibly attaining ∞ at some points), $\int_E f \, \mathrm{d}\mu \equiv \sup\left\{\int_E s \, \mathrm{d}\mu : s \leq f, s \text{ simple}\right\}$; 4) For any measurable function f (possibly attaining $\pm\infty$ at some points), $\int_E f \, \mathrm{d}\mu \equiv \int_E f^+ \, \mathrm{d}\mu - \int_E f^- \, \mathrm{d}\mu$, where $f^\pm \equiv \max(\pm f, 0)$, provided $\int_E |f| \, \mathrm{d}\mu = \int_E (f^+ + f^-) \, \mathrm{d}\mu < \infty$ (f is then said to be *Lebesgue integrable*). A function $f : E_x \mapsto E_y$ is *measurable* if $f^{-1}(\mathcal{F}(E_y)) \subseteq \mathcal{F}(E_x)$, where $(E_x, \mathcal{F}(E_x))$ and $(E_y, \mathcal{F}(E_y))$ are two measurable spaces. The generalized Lebesgue integral extends this concept of Lebesgue integral to measure spaces with generalized measures μ, not necessarily being Lebesgue measures (e.g., Hausdorff measures).

Define also the function $\eta(\xi) : \xi \in I \mapsto \bar{\mathbb{R}}^+$ such that for every value $\xi \in I$ it returns the μ_2-measure of the set $M_\xi^{\leq} \subset M$:

$$\eta(\xi) \equiv \mu_2\left(M_\xi^{\leq}\right) \equiv \int_{M_\xi^{\leq}} \mathrm{d}\mu_2 = \int_M \mathbf{1}_{\left[M_\xi^{\leq}\right]} \mathrm{d}\mu_2, \tag{D.2}$$

where the function $\mathbf{1}_{\left[M_\xi^{\leq}\right]}$ is the characteristic function on $M_\xi^{\leq}$. Let $\delta\eta(\xi, \mathrm{d}\xi)$ be the difference between the values of η at $\xi + \mathrm{d}\xi$ and ξ:

$$\begin{aligned}
\delta\eta(\xi, \mathrm{d}\xi) &\equiv \eta(\xi + \mathrm{d}\xi) - \eta(\xi) \equiv \mu_2\left(M_{\xi+\mathrm{d}\xi}^{\leq}\right) - \mu_2\left(M_\xi^{\leq}\right) \equiv & \text{(D.3)}\\
&\equiv \left(\int_{M_{\xi+\mathrm{d}\xi}^{\leq}} - \int_{M_\xi^{\leq}}\right) \mathrm{d}\mu_2 = \int_M \left(\mathbf{1}_{\left[M_{\xi+\mathrm{d}\xi}^{\leq}\right]} - \mathbf{1}_{\left[M_\xi^{\leq}\right]}\right) \mathrm{d}\mu_2 & \text{(D.4)}
\end{aligned}$$

and define formally $\mathrm{d}\eta(\xi)/\mathrm{d}\xi$ as the limit of $\delta\eta(\xi, \mathrm{d}\xi)/\mathrm{d}\xi$ when $\mathrm{d}\xi \to 0$:

$$\begin{aligned}
\frac{\mathrm{d}\eta(\xi)}{\mathrm{d}\xi} &\equiv \lim_{\mathrm{d}\xi\to 0} \frac{\delta\eta(\xi, \mathrm{d}\xi)}{\mathrm{d}\xi} = & \text{(D.5)}\\
&= \lim_{\mathrm{d}\xi\to 0} \frac{\left(\int_{M_{\xi+\mathrm{d}\xi}^{\leq}} - \int_{M_\xi^{\leq}}\right) \mathrm{d}\mu_2}{\mathrm{d}\xi} = & \text{(D.6)}\\
&= \lim_{\mathrm{d}\xi\to 0} \frac{\int_M \left(\mathbf{1}_{\left[M_{\xi+\mathrm{d}\xi}^{\leq}\right]} - \mathbf{1}_{\left[M_\xi^{\leq}\right]}\right) \mathrm{d}\mu_2}{\mathrm{d}\xi} = & \text{(D.7)}\\
&= \lim_{\mathrm{d}\xi\to 0} \int_M \frac{\mathbf{1}_{\left[M_{\xi+\mathrm{d}\xi}^{\leq}\right]} - \mathbf{1}_{\left[M_\xi^{\leq}\right]}}{\mathrm{d}\xi} \mathrm{d}\mu_2 = & \text{(D.8)}\\
&= \int_M \lim_{\mathrm{d}\xi\to 0} \left(\frac{\mathbf{1}_{\left[M_{\xi+\mathrm{d}\xi}^{\leq}\right]} - \mathbf{1}_{\left[M_\xi^{\leq}\right]}}{\mathrm{d}\xi}\right) \mathrm{d}\mu_2. & \text{(D.9)}
\end{aligned}$$

Define the generalized function:

$$\delta_{\left[M_\xi^{=}\right]} \equiv \lim_{\mathrm{d}\xi\to 0} \left(\frac{\mathbf{1}_{\left[M_{\xi+\mathrm{d}\xi}^{\leq}\right]} - \mathbf{1}_{\left[M_\xi^{\leq}\right]}}{\mathrm{d}\xi}\right) \equiv \frac{\mathrm{d}\mathbf{1}_{\left[M_\xi^{=}\right]}}{\mathrm{d}\xi}. \tag{D.10}$$

It can be considered as an operator such that, when applied to a function $f(P, \xi)$ defined on M, it returns the variation of $f(P, \xi)$ in the direction normal to the tangent space of $M_\xi^{=}$ on M at each

point $P \in M_{\xi}^{=}$. Then, equation D.9 results:

$$\frac{\mathrm{d}\eta(\xi)}{\mathrm{d}\xi} = \int_M \delta_{\left[M_{\xi}^{=}\right]} \mathrm{d}\mu_2. \tag{D.11}$$

Consider, in particular, μ_i, $i = 1, 2$, to be the α_i-dimensional *Hausdorff measure*[4], $\mathcal{H}^{\alpha_i}$ on $\mathbb{R}^n$, such that $\alpha_2 > \alpha_1$, and $\mathrm{d}\mu_2 = \mathrm{d}\mu_1 \mathrm{d}(\mu_2/\mu_1)$, where μ_2/μ_1 is the quotient of μ_2 by μ_1. Then, for a regular and smooth[5] manifold M:

$$\frac{d\eta(\xi)}{\mathrm{d}\xi} = \int_M \delta_{\left[M_{\xi}^{=}\right]} \mathrm{d}\mu_2 = \int_M \left[\delta_{\left[M_{\xi}^{=}\right]} d(\mu_2/\mu_1)\right] \mathrm{d}\mu_1 \equiv f(\xi) \int_{M_{\xi}^{=}} \mathrm{d}\mu_1 \tag{D.12}$$

$$= f(\xi)\mu_1(M_{\xi}^{=}) \equiv f(\xi)\Psi(\xi), \tag{D.13}$$

where the function $f(\xi)$ is defined according the Mean-Value Theorem (applicable since the manifold is regular and smooth):

$$f(\xi) \equiv \frac{\int_M \left[\delta_{\left[M_{\xi}^{=}\right]} d(\mu_2/\mu_1)\right] \mathrm{d}\mu_1}{\int_{M_{\xi}^{=}} \mathrm{d}\mu_1} \equiv \frac{\overline{\mathrm{d}[\mu_2/\mu_1|_{M_{\xi}^{=}}}}{\mathrm{d}\xi}. \tag{D.14}$$

Considering the explanation of the character of $\delta_{\left[M_{\xi}^{=}\right]}$, the function $f(\xi)$ can be interpreted as the average value of the variation with ξ of μ_2/μ_1 on the set $M_{\xi}^{=}$ (expressed as $\overline{\mathrm{d}[\mu_2/\mu_1|_{M_{\xi}^{=}}/\mathrm{d}\xi}$). Therefore, in order to measure sets $M_{\xi}^{=} \subset M \subset \mathbb{R}^n$ in the μ_1 Hausdorff measure, $\mu_1 \equiv \mathcal{H}^{\alpha_1}$, it is possible to use alternatively the μ_2 Hausdorff measure, $\mu_2 \equiv \mathcal{H}^{\alpha_2}$, on the set $M_{\xi,\mathrm{d}\xi} = M_{\xi+\mathrm{d}\xi}^{\leq} \cap M_{\xi}^{\leq} =$

[4] Let (E, d) be a metric space (with a distance d defined on the set E). The α-dimensional *Hausdorff measure* of the set $A \subset E$, $\mathcal{H}^{\alpha}(A) \in [0, +\infty]$, is defined as $\mathcal{H}^{\alpha}(A) \equiv \lim_{\delta \to 0^+} \mathcal{H}_{\delta}^{\alpha}(A)$, being $\mathcal{H}_{\delta}^{\alpha}(A) \equiv \inf\{\sum_{j=0}^{\infty} \omega_{\alpha} \left(\mathrm{diam}(B_j^{\delta})/2\right)^{\alpha} : B_j^{\delta} \subset E,\ \bigcup_{j=0}^{\infty} B_j^{\delta} \supset A,\ \mathrm{diam}(B_j^{\delta}) \leq \delta, \forall j = 0, 1, \ldots\}$, where $\mathrm{diam}(B_j^{\delta}) \equiv \sup_{x,y \in B_j^{\delta}} d(x, y)$, $\omega_{\alpha} = \pi^{\alpha/2}/\Gamma(\alpha/2+1)$, ($\Gamma(x)$ is the Gamma Function), $\delta > 0$, $\alpha \geq 0$ and the infimum is taken over all possible enumerable families of sets $\{B_0^{\delta}, B_1^{\delta}, \ldots, B_j^{\delta}, \ldots\}$ which are sufficiently small ($\mathrm{diam}(B_j) \leq \delta$) and which cover A. The limit exists since the function $\mathcal{H}_{\delta}^{\alpha}(E)$ is decreasing in δ: $\delta' < \delta \Rightarrow \bigcup_{i=0}^{\infty} B_i^{\delta'} \subset \bigcup_{j=0}^{\infty} B_j^{\delta} \Rightarrow \mathcal{H}_{\delta'}^{\alpha}(E) < \mathcal{H}_{\delta}^{\alpha}(E)$. The Hausdorff measure is a Borel external measure on $\mathbb{R}^n$ that generalizes the concept of length, area, and volume of sets in $\mathbb{R}^n$. For the particular case of a m-dimensional regular manifold $M \subset \mathbb{R}^n$, $\mathcal{H}^m(M)$ is the m-dimensional area of M. For $m = n$, $\mathcal{H}^n$ is the Lebesgue measure on $\mathbb{R}^n$. But as an external measure, $\mathcal{H}^{\alpha}$ is defined on every subset of $\mathbb{R}^n$, not only on regular manifolds.

[5] A smooth manifold is infinitely differentiable. In particular, a two-dimensional surface parametrized by variables (u, v) is smooth if the tangent vectors in the u and v directions satisfy: $\boldsymbol{t}_u \wedge \boldsymbol{t}_v \neq 0$.

$\{P \in M \mid \xi \leq \xi_M(P) < \xi + \mathrm{d}\xi\}$ divided by $f(\xi)\,\mathrm{d}\xi$ and then take the limit $\mathrm{d}\xi \to 0$:

$$\mu_1(M_{\xi}^{=}) = \lim_{\mathrm{d}\xi \to 0} \frac{\mu_2(M_{\xi,\mathrm{d}\xi})}{f(\xi)\mathrm{d}\xi}, \tag{D.15}$$

derived from equations D.6, D.13, and the relation $\mu_2(M_{\xi+\mathrm{d}\xi}^{\leq} \cap M_{\xi}^{\leq}) = \mu_2(M_{\xi+\mathrm{d}\xi}^{\leq}) - \mu_2(M_{\xi}^{\leq})$.

The *Hausdorff dimension*[6], α, of the sets $M_{\xi}^{=}, M_{\xi,\mathrm{d}\xi} \subset M$, satisfies:

$$\Delta\alpha \equiv \alpha(M_{\xi,\mathrm{d}\xi}) - \alpha(M_{\xi}^{=}) \geq 0. \tag{D.16}$$

Thus, equation D.15 implicitly indicates a reduction in the Hausdorff dimension of $\mu_2(M_{\xi,\mathrm{d}\xi})$ by taking the limit of it after dividing by $f(\xi)\,\mathrm{d}\xi$, obtaining $\mu_1(M_{\xi}^{=})$ as a result.

Consider the function $\Psi(\xi)$, introduced above, with the particular choice of the measure μ_1 as being a α-dimensional Hausdorff measure, $\mathcal{H}^{\alpha}$. Also, consider the density function $\tilde{\Psi}(\xi) \equiv f(\xi)\,\Psi(\xi) : \xi \in I \mapsto \mathbb{R}^+$ with the choice of μ_1 and μ_2 as Hausdorff measures of dimension α and $\alpha + \Delta\alpha$, respectively, $\mu_1 = \mathcal{H}^{\alpha}$ and $\mu_2 = \mathcal{H}^{\alpha+\Delta\alpha}$. From the definition of $\tilde{\Psi}(\xi)$ and $\eta(\xi)$, (equations D.1, D.2), and the relation between them given by equation D.13 it results:

$$\int_{\xi_{\min}}^{\xi_{\max}} \tilde{\Psi}(\xi)\,\mathrm{d}\xi = \int_{\xi_{\min}}^{\xi_{\max}} f(\xi)\,\Psi(\xi)\,\mathrm{d}\xi = \int_{\xi_{\min}}^{\xi_{\max}} \frac{\mathrm{d}\eta(\xi)}{\mathrm{d}\xi}\,\mathrm{d}\xi = \int_{\xi_{\min}}^{\xi_{\max}} \mathrm{d}\eta(\xi) = \tag{D.17}$$

$$= \eta(\xi_{\max}) - \eta(\xi_{\min}) = \mu_2(M) = \mathcal{H}^{\alpha+\Delta\alpha}(M). \tag{D.18}$$

Therefore, the integral of the density function $\tilde{\Psi}(\xi)$ of M over the range I of ξ is the $(\alpha + \Delta\alpha)$-dimensional Hausdorff measure of M. It can be normalized to obtain the corresponding probability

[6] The *Hausdorff dimension*, $\alpha(A) \geq 0$, of a subset A of a metric space (E, d), is defined as $\alpha(A) = \inf\{D \mid \lim_{r\to 0}[H_r^D(A)]\}$ being $H_r^D(A) = \inf \sum_{i\in I} \left(\mathrm{diam}(B_i^r)\right)^D$ where $\{B_i^r, i \in I, I \text{ countable set}\}$ is a countable r-cover of A and the infimum in H_r^D is over all countable r-covers of A. If A is a subset of $\mathbb{R}^n$ with any restricted norm-induced metric, this definition is equivalent to $\alpha(A) = -\lim_{r\to 0}[\log N_A(r)/\log r]$, where $N_A(r)$ is the minimum number of balls of radius r required to cover A. For a fixed set $A \subset E$ there exists at most one value α such that the α-dimensional Hausdorff measure of A, $\mathcal{H}^{\alpha}(A)$ is finite and positive. For $\alpha' > \alpha$, $\mathcal{H}^{\alpha'}(A) = 0$, whereas for $\alpha' < \alpha$, $\mathcal{H}^{\alpha'}(A) \to +\infty$. This result can be used equivalently to define the dimension of a set A, $\alpha(A)$ as the value for which its associated α-dimensional Hausdorff measure, $\mathcal{H}^{\alpha}(A)$, is finite and positive. For example, the Hausdorff dimension of a regular two-dimensional surface $M \subset \mathbb{R}^n$ is two, and $\mathcal{H}^2(M)$ (which coincides with the area of the surface) will be finite and positive, while $\mathcal{H}^1(M)$ (length of M) will be infinite, and $\mathcal{H}^3(M)$ (volume of M) will be zero. The Hausdorff dimensions of a set need not be integer (e.g., most fractals have a non-integer Hausdorff dimension).

density function:

$$\mathcal{P}(\xi) \equiv \frac{\tilde{\Psi}(\xi)}{\mathcal{H}^{\alpha+\Delta\alpha}(M)}, \qquad \text{with} \qquad \int_{\xi_{\min}}^{\xi_{\max}} \mathcal{P}(\xi)\,\mathrm{d}\xi = 1. \tag{D.19}$$

Depending on the distribution of the local property ξ throughout the manifold M, it will be appropriate to choose particular values of α and $\Delta\alpha$ for measuring the sets $M_{\xi}^{=}$ and $M_{\xi,\mathrm{d}\xi}$ in order to obtain a relevant $\tilde{\Psi}(\xi)$. For example, for a surface M in a three-dimensional euclidean space:

- If ξ is distributed mainly in patches of constant ξ, then a dimension $\alpha = 2$ with $\Delta\alpha = 0$ ($\Rightarrow \mu_1 = \mu_2$) would be appropriate: $\tilde{\Psi}(\xi)$ would then give the area of those patches for the particular values of ξ at which they appear (see Figure D.1). The sum of all those values would be the area of M ($\sum_i \tilde{\Psi}(\xi_i) = \mathcal{H}^2(M)$). By using this measure, subsets of Hausdorff dimension less than two (curves of constant ξ or isolated points of constant ξ) would not be reflected in $\tilde{\Psi}(\xi)$, since their associated $\mathcal{H}^2$ measure is null.

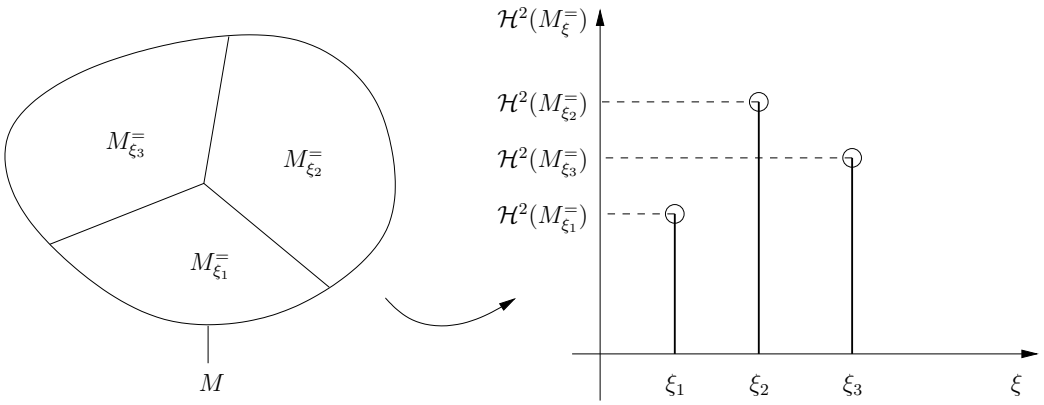


Figure D.1: $\Psi(\xi) = \tilde{\Psi}(\xi)$ (right) with $\alpha = 2$ for a surface M (left) with the local property ξ distributed in patches of constant ξ. Each point of that function (right) represents the area (two-dimensional Hausdorff measure) of the associated patch. Their discrete sum equals the total area of the surface M

- If ξ is smoothly distributed throughout M, the appropriate dimension to use is $\alpha = 1$ ($\mu_1 \equiv \mathcal{H}^1$), with $\Delta\alpha = 1$ since the sets $M_{\xi}^{=}$ will be curves of constant ξ (unitary Hausdorff dimension) or isolated points (null Hausdorff dimension). $\tilde{\Psi}(\xi)$ will be continuous and its integral with respect to ξ will be $\mathcal{H}^2(M)$ (according to equation D.18), that is, the area of M. If ξ is piece-wise smoothly distributed throughout M, that is, smooth except in the boundaries of patches of M with constant ξ (see Figure D.2), these patches will be reflected in $\Psi(\xi)$ as delta functions

at the corresponding value of ξ associated with each patch, such that the integral with respect to ξ equals (in the limit $\mathrm{d}\xi \to 0$) the area of the patch. The shape of $\tilde{\Psi}(\xi)$ (see Figure D.3)

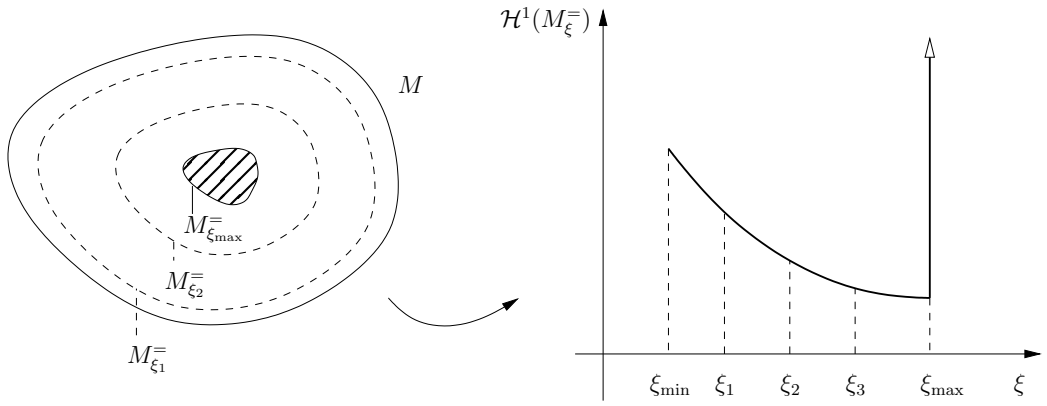


Figure D.2: $\Psi(\xi)$ (right) with $\alpha = 1$ for a surface M with ξ smoothly distributed throughout M (except one patch of constant $\xi = \xi_{\max}$). In the left diagram, dashed lines represent line contours of constant $\xi = \xi_1, \xi_2$, which have an associated finite value of $\Psi(\xi)$ (since their Hausdorff dimension equals the dimension of the measure used to obtain $\Psi(\xi)$, $\alpha(M^=_{\xi_{1,2}}) = 1$), whereas the central patch (filled with oblique lines pattern) of constant $\xi = \xi_{\max}$ has an associated value $\mathcal{H}^1(M^=_{\xi_{\max}}) \to \infty$, since its Hausdorff dimension is $\alpha(M^=_{\xi_{\max}}) = 2$

will depend on the function $f(\xi)$, that represents how 'distant' two different sets (curves, in general), $M^=_\xi$ and $M^=_{\xi+\mathrm{d}\xi}$, are. That distance, for each point $P \in M^=_\xi$, is measured along the coordinate n of the tangent plane at P normal to the arc length s of $M^=_\xi$, and then averaged over the whole set, thus resulting in a function of ξ only. Large values of $f(\xi)$ indicate that the property ξ varies slowly along n in average, whereas small values of $f(\xi)$ correspond to a rapid averaged variation of ξ with n.

The resulting $\tilde{\Psi}(\xi)$ (and, alternatively, $\mathcal{P}(\xi)$) of M can be regarded as a non-local characterization of the distribution of ξ throughout M.

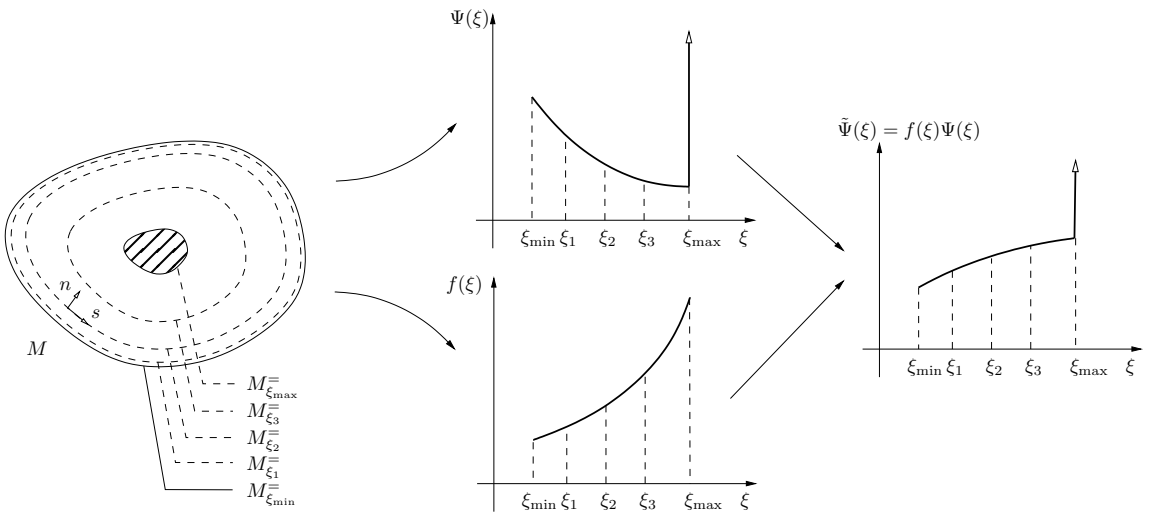


Figure D.3: $\tilde{\Psi}(\xi)$ (right) for a surface M with ξ smoothly distributed throughout M (except one patch of constant $\xi = \xi_{\max}$). In the left diagram, dashed lines represent line contours of constant $\xi = \xi_{\min}, \xi_1, \xi_2, \xi_3, \xi_{\max}$, such that they differ in a constant $\Delta\xi$. The central part of the figure represents $\Psi(\xi)$ (top) and $f(\xi)$ (bottom). $f(\xi)$ increases with ξ since the averaged distance between contours (normal to them along n) increases with ξ. On the right, the multiplication of both gives $\tilde{\Psi}(\xi)$

D.1 Conditions for existence of an explicit analytical solution

Consider an explicit parametrization of the surface M in terms of two parameters (u, v), and also an explicit parametrization of the local property ξ on M in terms of the same two parameters:

$$M : (u, v) \in (I_u, I_v) \mapsto \mathbb{R}^n, \tag{D.20}$$

$$\xi : (u, v) \in (I_u, I_v) \mapsto \mathbb{R}. \tag{D.21}$$

By choosing μ_2 to be the two-dimensional Hausdorff measure, the integrals in the function $\delta\eta$ defined above can be expressed in terms of the parametrization as:

$$\delta\eta(\xi, \mathrm{d}\xi) = \int_M \left(\mathbf{1}_{[M^{\leq}_{\xi+\mathrm{d}\xi}]} - \mathbf{1}_{[M^{\leq}_{\xi}]}\right)\mathrm{d}\mu_2 = \int_{(I_u, I_v)} \left(\mathbf{1}_{[M^{\leq}_{\xi+\mathrm{d}\xi}]} - \mathbf{1}_{[M^{\leq}_{\xi}]}\right)\theta(u, v)\,\mathrm{d}u\,\mathrm{d}v, \tag{D.22}$$

where $\mathrm{d}\mu_2 = \theta(u, v)\,\mathrm{d}u\,\mathrm{d}v$, and $\theta(u, v)$ depends on the parametrization of the surface.

Under the following constraints imposed on the parametrization, the function $\tilde{\Psi}(\xi) = \lim_{\mathrm{d}\xi \to 0}(\delta\eta(\xi, \mathrm{d}\xi)/\mathrm{d}\xi)$ can be obtained explicitly in terms of the parameters (u, v), providing interesting analytical solutions

of $\tilde{\Psi}(\xi)$ for certain surfaces:

1. If the functions $\xi(u,v)$ and $\theta(u,v)$ are both independent of one (the same one) of the two parameters (u,v) (the parameter v has been chosen for that purpose in this development without loss of generality), $\xi(u,v) \equiv \xi(u)$, $\theta(u,v) \equiv \theta(u)$, then $\tilde{\Psi}(\xi)$ can be written as:

$$\tilde{\Psi}(\xi(u)) = \lim_{\mathrm{d}\xi \to 0} \left(\int_{I_v} \mathrm{d}v \right) \left(\int_{I_u} \left(\mathbf{1}_{\left[M^{\leq}_{\xi+\mathrm{d}\xi}\right]} - \mathbf{1}_{\left[M^{\leq}_{\xi}\right]} \right) \theta(u)\, \mathrm{d}u \right) \tag{D.23}$$

$$= \Delta v \, \frac{\theta(u)\, \mathrm{d}u}{\mathrm{d}\xi} = \Delta v \, \frac{\theta(u)\, \mathrm{d}u}{\dfrac{\mathrm{d}\xi}{\mathrm{d}u}\, \mathrm{d}u} = \Delta v \frac{\theta(u)}{\dfrac{\mathrm{d}\xi}{\mathrm{d}u}(u)}, \tag{D.24}$$

   where $\Delta v = \int_{I_v} \mathrm{d}v$ is a constant.

2. If the map $\xi(u) : u \in I_u \subset \mathbb{R} \mapsto \xi \in I_\xi \subset \mathbb{R}$ is invertible (i.e., bijective[7]), there exists the inverse map $u = u(\xi) : \xi \in I_\xi \subset \mathbb{R} \mapsto u \in I_u \subset \mathbb{R}$ and equation D.24 can be finally written as an explicit analytical result:

$$\tilde{\Psi}(\xi) = \Delta v \frac{\theta(u(\xi))}{\dfrac{\mathrm{d}\xi}{\mathrm{d}u}(u(\xi))}. \tag{D.25}$$

   Note that the invertibility condition on the map $\xi(u)$ implies[8] that the first derivative $\mathrm{d}\xi/\mathrm{d}u$ exists and is non-zero $\forall u \in I_u$. Therefore, the function $\tilde{\Psi}(\xi)$ (that has $\mathrm{d}\xi/\mathrm{d}u$ in the denominator) is defined $\forall \xi \in I_\xi$.

   This invertibility condition can be relaxed still obtaining explicit analytical solution in those cases (see Figure D.4) in which there exists a countable number of **l**ocal **e**xtrema, $S_{le} = \{u^*_{le,p} \in I_u; p = 1, \ldots, N_{le}\}$, and a countable number of (surjective) subintervals, $S_I^{\mathrm{surj}} = \{I^{\mathrm{surj}}_{u,q} = [u^*_{\min,q}, u^*_{\max,q}] \subset I_u; q = 1, \ldots, N_I^{\mathrm{surj}}\}$ (the associated set of **e**xtreme **p**oints of those subintervals is called $S_{ep} = \{(u^*_{\min,q}, u^*_{\max,q}); q = 1, \ldots, N_{le}\}$), where the first derivative $\mathrm{d}\xi/\mathrm{d}u$ is null (i.e., $\mathrm{d}\xi/\mathrm{d}u \mid_{u^*} = 0$, $u^* \in S_{le} \cup S_I^{\mathrm{surj}}$). Define the set of points $S_P = \{u^*_j \in \{S_{le} \cup S_{ep} \cup \{u_{\min}, u_{\max}\}\}; j = 1, \ldots, (N_{le} + 2N_I^{\mathrm{surj}})\}$ ordered such that $u^*_j < u^*_{j+1}$. Define also

[7] A map $f : a \in A \leftrightarrow b \in B$ is *bijective* ($\forall a \in A \; \exists! \; b \in B \mid b = f(a)$) if it is *injective* ($\forall a \in A \; \exists \; b \in B \mid b = f(a)$) and *surjective* ($\forall b \in B \; \exists \; a \in A \mid b = f(a)$).

[8] The *inverse function theorem* states that a continuous function $f : x \in I_x \subset \mathbb{R} \mapsto y \in I_y \mathbb{R}$ is (locally) invertible (at $x' \in I_x$) if its first derivative is non-null, $df/dx \neq 0$ (at x'), that is, if f is strictly monotonic (at x'). f is invertible in I_x if it is locally invertible $\forall x \in I_x$

the set of (bijective) subintervals $S_I^{\text{bij}} = \{I_{u,r}^{\text{bij}} =]u_j^*, u_k^*[, r = 1, \ldots, N_I^{\text{bij}}; u_j^*, u_k^* \in S_P\}$ where there exists an invertible map $u_{I_{u,r}^{\text{bij}}}(\xi)$ (inverse function of $\xi(u)$ in the interval I_u^j). Note that $S_I^{\text{surj}} \cap S_I^{\text{bij}} = \emptyset$, and $I_u = S_I^{\text{bij}} \cup S_I^{\text{surj}} \cup S_{le}$. In that case, and assuming that there exists an explicit analytical expression for $u_j^* \in S_P$ in terms of ξ (which depends on the solvability of the equation $(\mathrm{d}\xi/\mathrm{d}u)(u) = 0$), then the function $\tilde{\Psi}(\xi)$ can still be explicitly obtained by the following analytical expression:

$$\tilde{\Psi}(\xi) = \Delta v \left[\sum_{r=1}^{N_I^{\text{bij}}} \frac{\theta(u_{I_{u,r}^{\text{bij}}}(\xi))}{\frac{\mathrm{d}\xi}{\mathrm{d}u}(u_{I_{u,r}^{\text{bij}}}(\xi))} + \sum_{p=1}^{N_{le}} \theta(u_{le,p}^*)\, \delta_0(\xi(u_{le,p}^*)) + \sum_{q=1}^{N_I^{\text{surj}}} \left(\int_{I_{u,q}^{\text{surj}}} \theta(u) \mathrm{d}u \right) \delta_1(\xi(I_{u,q}^{\text{surj}})) \right], \tag{D.26}$$

where the generalized functions $\delta_0(\xi)$ and $\delta_1(\xi)$ are zero everywhere except at ξ, where their value is an infinite with null and unitary total integral, respectively. The subintervals $I_{u,q}^* \in S_I^{\text{surj}}$ correspond to patches of the surface with constant ξ, which have a Hausdorff dimension of two, and therefore their one-dimensional measure is an integrable infinite such that, when integrated, it results the area of the patch (i.e., $\left(\int_{I_{u,q}^*} \theta(u)\, du\right) \Delta v = \int_{I_v} \int_{I_{u,q}^*} \theta(u)\, \mathrm{d}u\, \mathrm{d}v$).

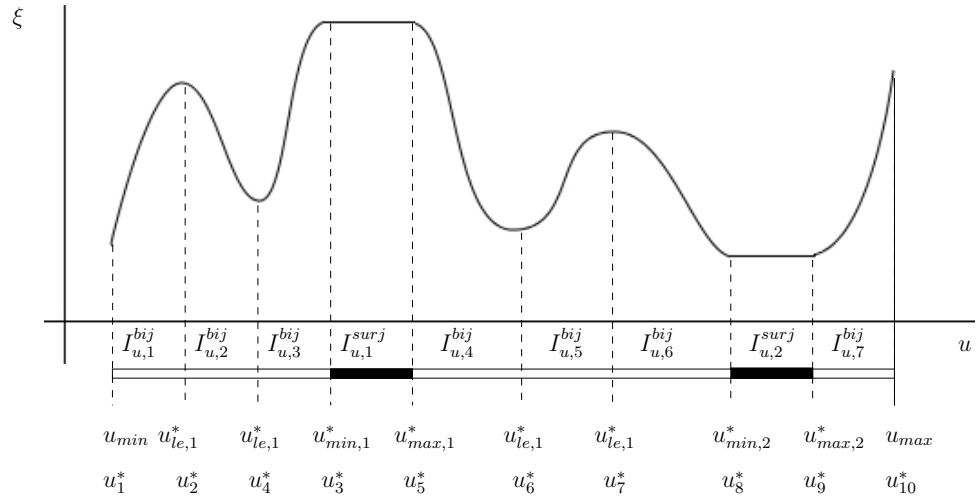


Figure D.4: Example of non-invertible $\xi(u)$ map

Common cases of existence of explicit analytical solution (complying with these two sufficient conditions) arise for cylindrical surfaces and surfaces of revolution, such that the property ξ preserves the cylindrical nature (being independent of the variable along the cylindrical axis) or the axisymmetric character of the surface (being independent of the azimuthal coordinate), and the

invertibility of the function relating ξ and the other variable of the parametrization is guaranteed, either globally or along subintervals.

D.2 Extension to multiple dimensions

A parallel development can be followed to define multi-variable density functions on a manifold M. For two local properties, ξ and ζ, we define $M^{=}_{\xi,\zeta}$ as the set of points P of M where $(\xi_M, \zeta_M)(P) = (\xi, \zeta)$, and $M^{=}_{\xi+\mathrm{d}\xi,\zeta+\mathrm{d}\zeta}$ as:

$$M^{=}_{\xi+\mathrm{d}\xi,\zeta+\mathrm{d}\zeta} = \left(M^{\leq}_{\xi+\mathrm{d}\xi} \cap M^{\leq}_{\xi}\right) \cap \left(M^{\leq}_{\zeta+\mathrm{d}\zeta} \cap M^{\leq}_{\zeta}\right) = \tag{D.27}$$

$$= \{P \in M \mid \xi \leq \xi_M(P) < \xi + \mathrm{d}\xi, \zeta \leq \zeta_M(P) < \zeta + \mathrm{d}\zeta\}\,. \tag{D.28}$$

$\Psi(\xi, \zeta)$ is now defined as $\Psi(\xi, \zeta) \equiv \mu_1\left(M^{=}_{\xi,\zeta}\right)$. Instead of $\mathrm{d}\eta(\xi)/\mathrm{d}\xi$, we have Jacobian determinant

$$J(\xi, \zeta) \equiv \left|\frac{\partial(\eta_\xi, \eta_\zeta)}{\partial(\xi, \zeta)}\right| = f(\xi, \zeta)\,\Psi(\xi, \zeta), \tag{D.29}$$

with $f(\xi, \zeta) \equiv \displaystyle\int_M \left[\delta_{\left[M^{=}_{\xi,\zeta}\right]} d(\mu_2/\mu_1)\right] \mathrm{d}\mu_1 \Big/ \int_{M^{=}_{\xi,\zeta}} \mathrm{d}\mu_1$. Then

$$\mu_1(M^{=}_{\xi,\zeta}) = \lim_{\mathrm{d}\xi,\mathrm{d}\zeta\to 0} \frac{\mu_2(M_{\xi,\zeta;\mathrm{d}\xi,\mathrm{d}\zeta})}{f(\xi, \zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta}. \tag{D.30}$$

We also define $\Delta\alpha \equiv \alpha(M_{\xi,\mathrm{d}\xi;\zeta,\mathrm{d}\zeta}) - \alpha(M_{\xi,\zeta})$ and $\tilde{\Psi}(\xi, \zeta) \equiv f(\xi, \zeta)\,\Psi(\xi, \zeta)$. Therefore:

$$\int_{\xi_{\min}}^{\xi_{\max}} \int_{\zeta_{\min}}^{\zeta_{\max}} \tilde{\Psi}(\xi, \zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta = \int_{\xi_{\min}}^{\xi_{\max}} \int_{\zeta_{\min}}^{\zeta_{\max}} f(\xi, \zeta)\,\Psi(\xi, \zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta = \tag{D.31}$$

$$= \int_{\xi_{\min}}^{\xi_{\max}} \int_{\zeta_{\min}}^{\zeta_{\max}} J(\xi, \zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta = \mu_2(M) = \mathcal{H}^{\alpha+\Delta\alpha}(M), \tag{D.32}$$

and the corresponding joint probability density function can be obtained by normalization as:

$$\mathcal{P}(\xi, \zeta) \equiv \frac{\tilde{\Psi}(\xi, \zeta)}{\mathcal{H}^{\alpha+\Delta\alpha}(M)}, \qquad \text{with} \qquad \int_{\xi_{\min}}^{\xi_{\max}} \int_{\zeta_{\min}}^{\zeta_{\max}} \mathcal{P}(\xi, \zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta = 1. \tag{D.33}$$

The one-dimensional (probability) density functions in terms of each variable can be directly obtained from the multi-dimensional one by integration with respect to the rest of variables. They are named *marginal* (probability) density functions. For the two-dimensional case, the corresponding marginal probability density functions are $\mathcal{P}_\xi(\xi) = \int_{\zeta_{\min}}^{\zeta_{\max}} \mathcal{P}(\xi,\zeta)\,\mathrm{d}\zeta$ and $\mathcal{P}_\zeta(\zeta) = \int_{\xi_{\min}}^{\xi_{\max}} \mathcal{P}(\xi,\zeta)\,\mathrm{d}\xi$.

An example of application is the use of the joint and/or marginal probability density functions of two differential-geometry properties of a surface M, such as the principal curvatures (κ_1, κ_2) or the shape index and curvedness (Υ, Λ), in terms of area-coverage on M, to provide a non-local geometrical characterization of such surface M.

Appendix E

Gauss-Bonnet theorem in the shape index, curvedness space

For any compact two-dimensional Riemann manifold without boundaries, M, the Gauss–Bonnet theorem states that the integral of the Gaussian curvature, K, over the manifold with respect to area, A, equals 2π times its *Euler characteristic*, χ:

$$\int_M K \mathrm{d}A = 2\pi\chi(M). \tag{E.1}$$

This formula relates the geometry of the surface (given by the integration of the Gaussian curvature, a differential-geometry property) to its topology (given by the Euler characteristic). The Euler characteristic of a surface is related to its *genus* [1] by $\chi = 2 - 2g$. From the relation among shape index, curvedness, and mean and Gaussian curvatures stated in Appendix C (see equations C.38 and C.39), the following relation can be obtained:

$$K = -\Lambda^2 \cos(\pi\Upsilon). \tag{E.2}$$

Then, the Gauss-Bonnet theorem can be restated in terms of the shape index and curvedness as

$$\int_M \Lambda^2 \cos(\pi\Upsilon)\, \mathrm{d}A = 4\pi[g(M) - 1]. \tag{E.3}$$

[1] The genus of an orientable surface is a topological invariant (as is the Euler characteristic) defined as the largest number of non-intersecting simple closed curves that can be drawn on the surface without disconnecting it.

Furthermore, considering the non-dimensionalization of the curvedness introduced in §2.2, $C = \mu\Lambda$ ($\mu \equiv 3V/A$, where V is the volume and A the area of the surface) and taking into account that cosine is a symmetric function and thus $\cos(\pi\Upsilon) = \cos(\pi|\Upsilon|) \equiv \cos(\pi S)$, then equation E.1 can be rewritten as

$$\int_M C^2 \cos(\pi S)\,\mathrm{d}A = 4\pi\mu^2[g(M) - 1]. \tag{E.4}$$

The left-hand side can be expressed in terms of the $\{S, C\}$ area-based joint probability density function of the surface, $\mathcal{P}(S, C)$:

$$\int_M C^2 \cos(\pi S)\,\mathrm{d}A = A \cdot \int\int C^2 \cos(\pi S)\,\mathcal{P}(S, C)\,\mathrm{d}S\,\mathrm{d}C. \tag{E.5}$$

Considering the stretching parameter, $\lambda \equiv \sqrt[3]{36\pi}(V^{2/3}/A)$, also introduced in §2.2, the Gauss-Bonnet theorem finally results in an integral relation between the $\{S, C\}$ area-based joint probability density function, $\mathcal{P}$, the stretching parameter, λ, and the genus of the surface, g:

$$\int\int C^2 \cos(\pi S)\,\mathcal{P}(S, C)\,\mathrm{d}S\,\mathrm{d}C = \lambda^3[g(M) - 1]. \tag{E.6}$$

Appendix F

Definition of feature center and upper and lower distances of a probability density function.

Consider a real-valued random variable X with probability density function $f(x)$, $x \in \mathbb{R}$. We define the *feature center* $\hat{x}$ as

$$\hat{x} \equiv \begin{cases} \bar{x} - d_l\sqrt{1-(d_l/d_u)^2} & \text{if } d_l < d_u \\ \bar{x} + d_u\sqrt{1-(d_u/d_l)^2} & \text{if } d_l > d_u \end{cases} \tag{F.1}$$

where $\bar{x}$ is the *mean* or *expected value* of X, $\bar{x} \equiv \int x f \mathrm{d}x$. The *lower* and *upper distances* are defined by

$$d_l \equiv \sqrt{\frac{\int_{x\leq\bar{x}} (\bar{x}-x)^2 f \mathrm{d}x}{\int_{x\leq\bar{x}} f \mathrm{d}x}}, \qquad d_u \equiv \sqrt{\frac{\int_{x\geq\bar{x}} (\bar{x}-x)^2 f \mathrm{d}x}{\int_{x\geq\bar{x}} f \mathrm{d}x}}. \tag{F.2}$$

The feature center can be interpreted as a correction to the mean that accounts for the asymmetry (skewness) of the density function $f(x)$ with respect to its mean, defining a new point closer to the region of higher density. When the probability density function $f(x)$ is symmetric, the feature center and mean coincide ($\hat{x} = \bar{x}$). The upper and lower distances, d_u and d_l, can be regarded as the r.m.s. of the part of the pdf above and below its mean value, respectively. A graphical example is shown in Figure F.1, for a probability density function $f(x) = x^2 \exp\left(-\sqrt{x}\right) / \int_0^\infty \xi^2 \exp\left(-\sqrt{\xi}\right) \mathrm{d}\xi$ that shows a long tail in one direction. The mean, $\bar{x}$, feature center, $\hat{x}$, and lower and upper distances, d_l

and d_u, are superimposed on the probability density function. These definitions can be immediately extended to higher-dimensional probability density functions.

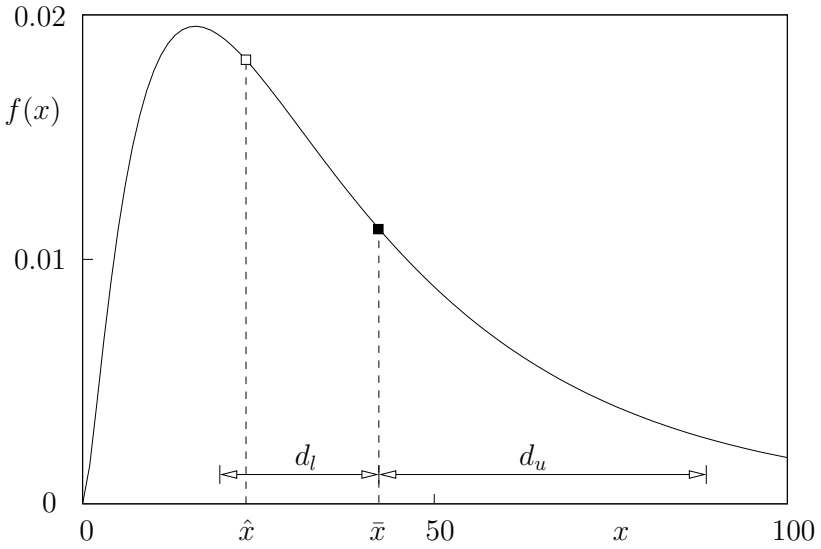
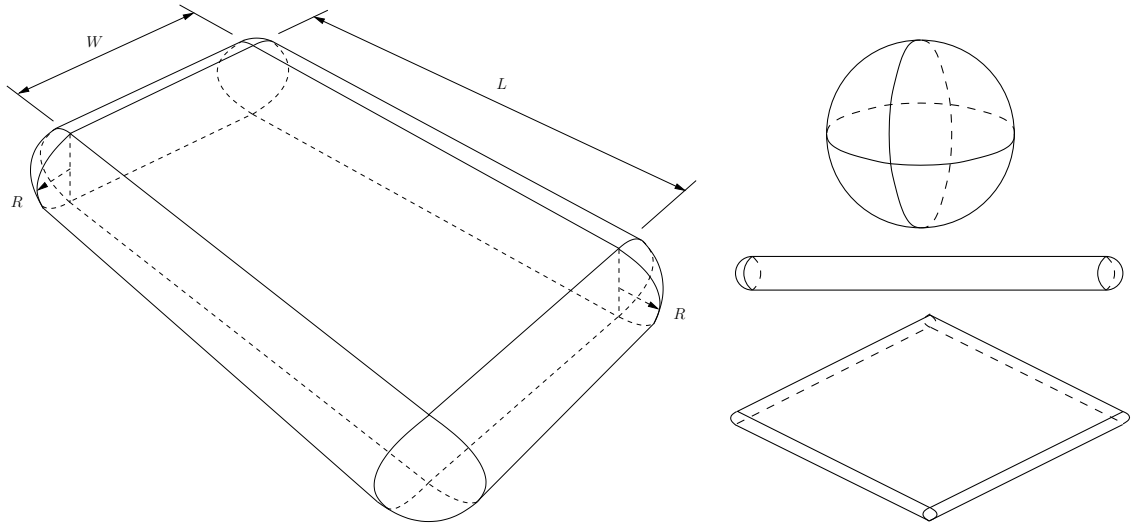


Figure F.1: Mean ($\bar{x}$) and feature ($\hat{x}$) centers and upper (d_u) and lower (d_l) distances for a sample asymmetric probability density function, $f(x)$

Appendix G

Analytic geometric characterization of limiting surfaces.

Consider the generic surface in Figure G.1(a). It consists of two planar parallel sheets of area $L\,W$ separated a distance of $2R$; four halves of circular cylinders of radii R and lengths L and W by pairs, tangent to the planar sheets that they connect; and four quarters of a sphere of radius R tangent to the circular cylinders. The resulting surface is closed. The surface is $\mathcal{C}^1$ along the curves of tangency among its parts (across which curvature is discontinuous) and $\mathcal{C}^2$ everywhere else. The area-based jpdf of S and C is thus still applicable.



(a) (b)

Figure G.1: Generic structure (a) and limiting cases (b)

Define $\xi \equiv L/R$, $\eta \equiv W/R$. Note that for $\xi = \eta = 0$ the surface is a sphere, for $\xi \gg 1$ and $\eta = 0$ (and vice versa) the surface is a circular tube with spherical caps (more stretched as ξ increases), and for $\xi, \eta \gg 1$ the surface is predominantly sheet-like (see Figure G.1(b)).

The area and volume of this surface are:

$$A = 4\pi R^2 \left[1 + \frac{1}{2}(\xi + \eta) + \frac{1}{2\pi}\xi\eta\right], \qquad V = \frac{4}{3}\pi R^3 \left[1 + \frac{3}{4}(\xi + \eta) + \frac{3}{2\pi}\xi\eta\right]. \tag{G.1}$$

Therefore

$$\mu \equiv \frac{3V}{A} = R\,\frac{1 + \frac{3}{4}(\xi + \eta) + \frac{3}{2\pi}\xi\eta}{1 + \frac{1}{2}(\xi + \eta) + \frac{1}{2\pi}\xi\eta}, \qquad \lambda \equiv \sqrt[3]{36\pi}\,\frac{V^{2/3}}{A} = \frac{\left[1 + \frac{3}{4}(\xi + \eta) + \frac{3}{2\pi}\xi\eta\right]^{2/3}}{1 + \frac{1}{2}(\xi + \eta) + \frac{1}{2\pi}\xi\eta}. \tag{G.2}$$

The principal curvatures, κ_1 and κ_2, are both $1/R$ in the spherical regions, $1/R$ and 0 respectively in the circular cylindrical regions, and both nil in the planar regions of such surface. Thus the dimensionless curvedness associated with each region is $C_{sph} = \mu/R$, $C_{cyl} = \mu/\sqrt{2}R$, $C_{pla} = 0$, respectively. The absolute value of the shape index is $S_{sph} = 1$ for the spherical regions and $S_{cyl} = 1/2$ for the circular cylindrical regions, while its value is undefined for the planar regions. For the purpose of this illustrative example, define such a value as $\gamma \in [0, 1]$.

The mean values of S and C for the surface, in terms of the dimensionless parameters ξ and η, result:

$$\bar{S} = \frac{1}{A}\left[S_{sph}A_{sph} + S_{cyl}A_{cyl} + S_{pla}A_{pla}\right] = \frac{1 + \frac{1}{4}(\xi + \eta) + \frac{1}{2\pi}\xi\eta\gamma}{1 + \frac{1}{2}(\xi + \eta) + \frac{1}{2\pi}\xi\eta}, \tag{G.3}$$

$$\bar{C} = \frac{1}{A}\left[C_{sph}A_{sph} + C_{cyl}A_{cyl} + C_{pla}A_{pla}\right] = \frac{\left[1 + \frac{3}{4}(\xi + \eta) + \frac{3}{2\pi}\xi\eta\right]\left[1 + \frac{1}{2\sqrt{2}}(\xi + \eta)\right]}{1 + \frac{1}{2}(\xi + \eta) + \frac{1}{2\pi}\xi\eta}. \tag{G.4}$$

In the limiting cases:

(i) for a sphere ($\xi = \eta = 0$): $\bar{S} = \bar{C} = 1$;

(ii) for a predominantly tube-like surface ($\xi \gg 1$, $\eta = 0$): $\bar{S} \approx 1/2$, $\bar{C} \approx 3/2\sqrt{2} \approx 1.06$;

(iii) for a predominantly sheet-like surface ($\xi = \eta \gg 1$): $\bar{S} \approx \gamma$, $\bar{C} \approx 0$.

Figure G.2 shows the dependence on ξ of $\bar{S}$, $\bar{C}$, and λ for the two last cases (surfaces becoming, as ξ increases, tube-like ($\eta = 0$) and sheet-like (with $\eta = \xi$ for simplicity)), starting from the sphere limit ($\xi = 0$). A particular value of γ has been chosen, without loss of generality, in order to represent the limit $\bar{S}_{sheet}$ graphically. In a general sheet-like surface, γ can take any value between 0 and 1, depending on its particular configuration. In the limiting cases ($\xi = 0$ and $\eta = 0$; $\xi \gg 1$ and $\eta = 0$; $\xi = \eta \gg 1$) $\hat{S} \approx \bar{S}$, $\hat{C} \approx \bar{C}$. Thus, a surface predominantly blob-, tube- or sheet-like can be distinguished based on its values of $\hat{S}, \hat{C}, \lambda$.

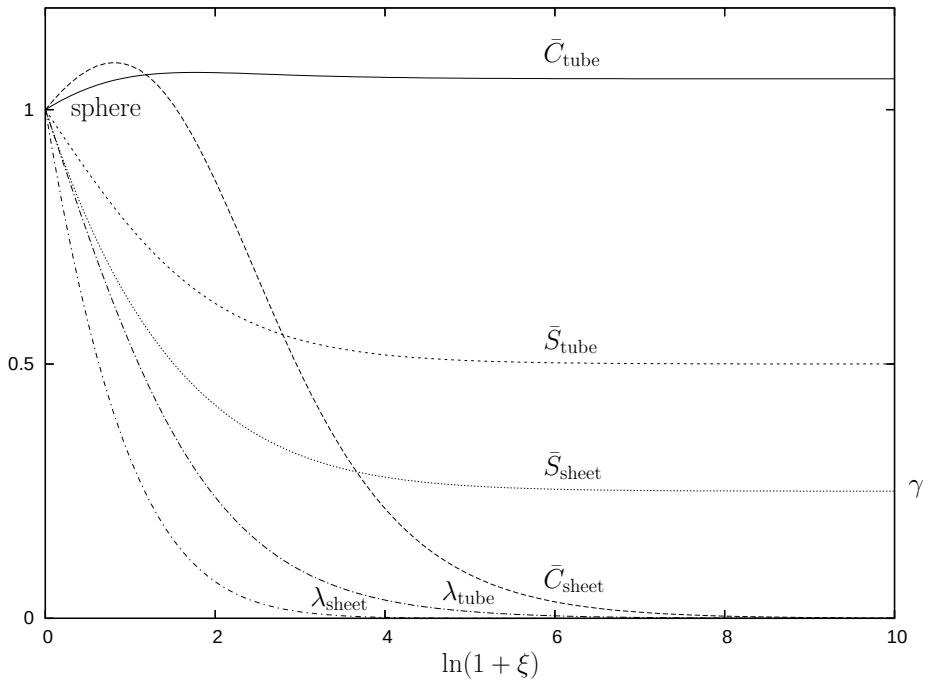


Figure G.2: $\bar{S}$, $\bar{C}$, and λ as a function of ξ for the tube-like and sheet-like limits, evolving from the sphere limit ($\xi = 0$). Note that the abscissa has been rescaled as $\ln(1+\xi)$ to show more clearly the transition region

Appendix H

Stratified random sampling with disproportionate allocation

When a multi-scale decomposition is applied to the scalar field from which the structures are educed, large differences among the number of structures obtained for each scale are to be expected. Larger scales will generally have a smaller number of educed structures than smaller scales. This difference in number can sometimes be of several orders of magnitude, particularly when analyzing fields with high grid resolution that results in a larger number of scales.

Thus, when structures of all scales are considered in the clustering algorithm, after the geometrical characterization, those structures (and their geometries) corresponding to the largest scale can be under-represented owing to the much smaller population they have compared to the others. In such a scenario, it can be beneficial to apply, prior to the clustering algorithm itself, a sampling of the population that takes into account the uneven sizes of the strata in which it can be divided.

We use a disproportionate stratification that considers the variance of the mutually exclusive strata to determine the sample size for each stratum. If n_o is the sample size of the stratum with the minimum standard deviation, $\sigma_o = \min\{\sigma_h, \forall h\}$, then the sample size, n_h, of any other stratum, h, with standard deviation σ_h will be proportional to $(\sigma_h/\sigma_o)\ n_o$. Therefore, those strata with higher variances will have also a higher number of elements to represent them in the clustering algorithm, accounting for their higher diversity. We take n_o as the population size of that stratum with the minimum standard deviation, N_o, since the purpose of this sampling is not to reduce the global population size, but to have a more balanced representation of the different groups present in it for

a better clustering.

After the disproportionate stratification, for those strata with $n_h < N_h$, where N_h is the population size of the stratum h, we take a random sample of n_h out of the N_h elements. Otherwise, the complete population is considered for that stratum.

Bibliography

Andreotti, B. 1997 Studying Burgers' models to investigate the physical meaning of the alignments statistically observed in turbulence. *Phys. Fluids* **9**, 735–742.

Antonia, R. A., Chambers, A. J., Friehe, C. A. & Van Atta, C. W. 1979 Temperature ramps in the atmospheric surface layer. *J. Atmos. Sci.* **36**, 99–108.

Ashurst, W. T., Kerstein, A. R., Kerr, R. M. & Gibson, C. H. 1987 Alignment of vorticity and scalar gradient with strain rate in simulated Navier-Stokes turbulence. *Phys. Fluids* **30**, 2343–2353.

Batchelor, G. K. 1959 Small-scale variation of convected quantities like temperature in turbulent fluid. Part 1. General discussion and the case of small conductivity. *J. Fluid Mech* **5**, 113–133.

Batchelor, G. K. & Townsend, A. A. 1949 The nature of turbulent motion at large wavenumbers. *Proc. R. Soc. Lond. A* **199**, 238–255.

Bellman, R. 1961 *Adaptive Control Processes: A Guided Tour*. Princeton University Press.

Berkhin, P. 2002 Survey of clustering data mining techniques. *Tech. Rep*. Accrue Software, San Jose, CA.

Bermejo-Moreno, I. & Pullin, D. I. 2008 On the non-local geometry of turbulence. *J. Fluid Mech.* **603**, 101–135.

Bradshaw, P. & Koh, Y. M. 1981 A note on Poisson's equation for pressure in a turbulent flow. *Phys. Fluids* **24**, 777.

Brasseur, J.G. & Wang, Q. 1992 Structural evolution of intermittency and anisotropy at different scales analyzed using three-dimensional wavelet transforms. *Phys. Fluids A* **4**, 2538–2554.

Brown, G. L. & Roshko, A. 1974 On density effects and large structure in turbulent mixing layers. *J. Fluid Mech.* **64**, 775–816.

Campbell, R. J. & Flynn, P. J. 2001 A survey of free-form object representation and recognition techniques. *Comput. Vis. Image Underst.* **81** (2), 166–210.

Candès, E., Demanet, L., Donoho, D. & Ying, L. 2005 Fast discrete curvelet transforms. *Multiscale Model. Simul.* **5**, 861–899.

Candès, E. J. & Donoho, D. L. 2003*a* Continuous curvelet transform: I Resolution of the wavefront set. *Appl. Comput. Harmon. Anal.* **19**, 162–197.

Candès, E. J. & Donoho, D. L. 2003*b* Continuous curvelet transform: II Discretization and frames. *Appl. Comput. Harmon. Anal.* **19**, 198–222.

Celani, A., Lanotte, A., Mazzino, A. & Vergassola, M. 2001 Fronts in passive scalar turbulence. *Phys. Fluids* **13**, 1768–1783.

Chen, S., Sreenivasan, K. R. & Nelkin, M. 1997 Inertial range scalings of dissipation and enstrophy in isotropic turbulence. *Phys. Rev. Letters* **79**, 1253–1256.

Chen, S. G. & Wu, J. Y. 2004 Estimating normal vectors and curvatures by centroid weights. *Comput. Aided Geom. Des.* **21** (5), 447–458.

Chen, X. & Schmitt, F. 1992 Intrinsic surface properties from surface triangulation. In *ECCV '92: Proc. Second European Conference on Computer Vision,*, 739–743. London, UK: Springer-Verlag.

Chong, M.S., Perry, A.E. & Cantwell, B.J. 1990 A general classification of three-dimensional flow field. *Phys. Fluids A* **2**, 765–783.

Corrsin, S. 1952 Heat transfer in isotropic turbulence. *J. Appl. Phys.* **23**, 113–118.

Dong, C. & Wang, G. 2005 Curvatures estimation on triangular mesh. *J. Zhejiang University SCIENCE* **6A** (I), 128–136.

Dorai, C. & Jain, A. K. 1997 Shape Spectrum Based View Grouping and Matching of 3D Free-Form Objects. *IEEE Trans. Pattern Anal. Mach. Intell.* **19** (10), 1139–1146.

Farge, M. 1992 Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.* **24**, 395–457.

Farge, M., Kevlahan, N., Perrier, V. & Goirand, E. 1996 Wavelets and turbulence. *IEEE Proceedings, Special Issue on Wavelets, ed. I Daubechies and J. Kovasevic* **84, 4**, 639–669.

Farge, M., Pellegrino, G. & Schneider, K. 2001 Coherent vortex extraction in 3D turbulent flows using orthogonal wavelets. *Phys. Rev. Lett.* **87** (5), 054501.

Farge, M., Schneider, K. & Kevlahan, N. 1999 Non-Gaussianity and Coherent Vortex Simulation for two-dimensional turbulence using an adaptative orthogonal wavelet basis. *Phys. Fluids* **11** (8), 2187–2201.

Farge, M., Schneider, K., Pellegrino, G., Wray, A. A. & Rogallo, R. S. 2003 Coherent vortex extraction in three-dimensional homogeneous turbulence: Comparison between CVS-wavelet and POD-Fourier decompositions. *Phys. Fluids* **15** (10), 2886–2896.

Hamlington, P. E., Schumacher, J. & Dahm, W. J. A. 2008 Local and nonlocal strain rate fields and vorticity alignment in turbulent flows. *Phys. Rev. E* **77**, 026303.

He, G., Chen, S., Kraichan, R. H., Zhang, R. & Zhou, Y. 1998 Statistics of dissipation and enstrophy induced by localized vortices. *Phys. Rev. Letters* **81**, 4639.

Horiuti, K. 2001 A classification method for vortex sheet and tube structures in turbulent flows. *Phys. Fluids A* **13** (12), 3756–3774.

Horiuti, K. & Fujisawa, T. 2008 The multi-mode stretched spiral vortex in homogeneous isotropic turbulence. *J. Fluid Mech.* **595**, 341–366.

Horiuti, K. & Takagi, Y. 2005 Identification method for vortex sheet structures in turbulent flows. *Phys. Fluids* **17** (12), 121703.

Hou, T.Y. & Li, R. 2006 Dynamic depletion of vortex stretching and non-blowup of the 3-D incompressible Euler equations. *J. Nonlinear Sci.* **16**, 639–664.

Hunt, J. C. R., Wray, A. A. & Moin, P. 1988 Eddies, stream, and convergence zones in turbulent flows. *Tech. Rep*. Center for Turbulence Research Report CTR-S88.

Ishihara, T., Kaneda, Y., Yokokawa, M., Itakura, K. & Uno, A. 2003 Spectra of energy dissipation, enstrophy and pressure by high-resolution direct numerical simulations of turbulence in a periodic box. *Journal of the Physical Society of Japan* **72**, 983–986.

Iyer, N., Jayantiand, S., Lou, K., Kalyanaraman, Y. & Ramani, K 2005 Three dimensional shape searching: state-of-the-art review and future trends. *Computer Aided Design* **37**, 509–530.

Jaberi, F. A., Miller, R. S., Madnia, C. K. & Givi, P. 1996 Non-Gaussian scalar statistics in homogeneous turbulence. *J. Fluid Mech.* **313**, 241–282.

Jayesh & Warhaft, Z. 1992 Probability distribution, conditional dissipation, and transport of passive temperature fluctuations in grid-generated turbulence. *Phys. Fluids* **4** (10), 2292–2307.

Jeong, J. & Hussain, F. 1995 On the identification of a vortex. *J. Fluid Mech.* **285**, 69–94.

Jiménez, J. 1992 Kinematic alignment effects in turbulent flows. *Phys. Fluids A* **4**, 652–654.

Jiménez, J., Wray, A. W., Saffman, P. G. & Rogallo, R. S. 1993 The structure of intense vorticity in isotropic turbulence. *J. Fluid Mech.* **255**, 65–90.

Kennedy, D. A. & Corrsin, S. 1961 Spectral flatness factor and intermittency in turbulence and in non-linear noise. *J. Fluid Mech.* **10**, 366–370.

Kerr, R. M. 1985 Higher-order derivative correlations and the alignment of small-scale structures in isotropic numerical turbulence. *J. Fluid Mech.* **153**, 31–58.

Koenderink, J. J. & van Doorn, A. J. 1992 Surface shape and curvature scales. *Image Vision Comput.* **10** (8), 557–565.

Kolmogorov, A. N. 1941*a* Dissipation of energy in a locally isotropic turbulence. *Dokl. Nauk. SSSR.* **32**, 16–18.

Kolmogorov, A. N. 1941*b* The local structure of turbulence in incompressible viscous fluids at very large Reynolds numbers. *Dokl. Nauk. SSSR.* **30**, 301–305.

Kolmogorov, A. N. 1962 A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds numbers. *J. Fluid Mech.* **13**, 82–85.

Kraichnan, R. H. 1974 On Kolmogorov's inertial-range theories. *J. Fluid Mech.* **62**, 305–330.

Landau, L. D. & Lifshitz, E. M. 1959 *Fluid Mechanics*. Pergamon.

Lundgren, T. S. 1982 Strained spiral vortex model for turbulent fine structure. *Phys. Fluids* **25**, 2193–2203.

Meneveau, C. 1991 Analysis of turbulence in the orthonormal wavelet representation. *J. Fluid Mech.* **232**, 469–520.

Meyer, M., Desbrun, M., Schroder, P. & Barr, A. H. 2003 Discrete differential operators for triangulated 2-manifolds. *Visualizations and Mathematics III,* 33–57.

Misra, A. & Pullin, D. I. 1997 A vortex-based subgrid stress model for large-eddy simulations. *Phys. Fluids* **9**, 2443–2454.

Miura, H. & Kida, S. 1997 Indentification of tubular vortices in turbulence. *J. Phys. Soc. Japan* **66** (5), 1331–1334.

Moisy, F. & Jiménez, J. 2004 Geometry and clustering of intense structures in isotropic turbulence. *J. Fluid Mech.* **513**, 111–133.

Ng, A., Jordan, M. & Weiss, Y. 2001 On spectral clustering: Analysis and an algorithm. *Adv. Neural Inform. Proc. Sys.* **14**.

Nomura, K. K. & Post, G. K. 1998 The structure and dynamics of vorticity and rate of strain in incompressible homogeneous turbulence. *J. Fluid Mech.* **377**, 65–97.

O'Gorman, P. A. & Pullin, D. I. 2004 On modal time correlations of turbulent velocity and scalar fields. *J. Turbulence* **5**, 035.

Okamoto, N., Yoshimatsu, K., Shneider, K., Farge, M. & Kaneda, Y. 2007 Coherent vortices in high resolution direct numerical simulation of homogeneous isotropic turbulence: A wavelet viewpoint. *Phys. Fluids* **19**, 115109.

Onsager, L. 1945 The distribution of energy in turbulence. *Phys. Rev.* **68**, 286.

Osada, R., Funkhouser, T., Chazelle, B. & Dobkin, D. 2001 Matching 3D models with shape distributions. In *SMI '01: Proc. Intl. Conf. on Shape Modeling & Applications*, 154–166. Washington, DC, USA: IEEE Computer Society.

Overholt, M. R. & Pope, S. B. 1996 Direct numerical simulation of a passive scalar with imposed mean gradient in isotropic turbulence. *Phys. Fluids* **8**, 3128–3148.

Perry, A. E. & Chong, M. S. 1987 A description of eddying motions and flow patterns using critical-point concepts. *Annu. Rev. Fluid Mech.* **19**, 125–55.

Pullin, D.I. & Saffman, P.G. 1993 On the Lundgren-Townsend model of turbulent fine scales. *Phys. Fluids A* **5**, 126–145.

Pullin, D. I. & Lundgren, T. S. 2001 Axial motion and scalar transport in stretched spiral vortices. *Phys. Fluids* **13**, 2553–2563.

Pullin, D. I. & Saffman, P. G. 1997 Vortex dynamics in turbulence. *Annu. Rev. Fluid Mech.* **30**, 31–51.

Richardson, L. F. 1922 *Weather Prediction by Numerical Process*. Cambridge University Press.

Rousseeuw, P. J. 1987 Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Maths* **20**, 53–65.

Ruetsch, G. R. & Maxey, M. R. 1992 The evolution of small-scale structures in isotropic homogeneous turbulence. *Phys. Fluids A* **4 (12)**, 2747–2760.

Schumacher, J. & Sreenivasan, K. R. 2005 Statistics and geometry of passive scalars in turbulence. *Phys. Fluids* **17**, 125107.

Schumacher, J., Sreenivasan, K. R. & Yeung, P. K. 2005 Very fine structures in scalar mixing. *J. Fluid Mech.* **531**, 113–122.

Schwarz, G. 1978 Estimating the dimension of a model. *Ann. Statist.* **6** (2), 461–464.

Shumacher, J. & Sreenivasan, K. R. 2005 Very fine structures in scalar mixing. *J. Fluid Mech.* **531**, 113–122.

Siggia, E. D. 1981 Numerical study of small-scale intermittency in three-dimensional turbulence. *J. Fluid Mech.* **107**, 375–406.

Sreenivasan, K. R. 2004 Possible effects of small-scale intermittency in turbulent reacting flows. *Flow, Turbulence and Combustion* **72**, 115–131.

Sreenivasan, K. R. & Meneveau, C. 1988 Singularities of the the equations of fluid motion. *Phys. Rev. A* **38**, 6287–6295.

Tanaka, M. & Kida, S. 1993 Characterization of vortex tubes and sheets. *Phys. Fluids A* **5**, 2079–2082.

Tao, B., Katz, J. & Meneveau, C. 2000 Geometry and scale relationships in high Reynolds number turbulence determined from three-dimensional holographic velocimetry. *Phys. Fluids* **12**, 941–944.

Taubin, G. 1995 Estimating the tensor of curvature of a surface from a polyhedral approximation. *ICCV '95: Proc. Fifth Intl. Conf on Computer Vision,* 902–907.

Tennekes, A. A. 1968 Simple model for the small-scale structure of turbulence. *Phys. Fluids* **11**, 669–671.

Tennekes, H. & Lumley, J. L. 1974 *A First Course in Turbulence*. The MIT Press.

Townsend, A. A. 1951 On the fine-scale structure of turbulence. *Proc. R. Soc. London Ser. A* **208**, 534–542.

Tsinober, A., Kit, E. & Dracos, T. 1992 Experimental investigation of the field of velocity gradients in turbulent flows. *J. Fluid Mech.* **242**, 169–192.

Vincent, A. & Meneguzzi, M. 1994 The dynamics of vorticity tubes in homogeneous turbulence. *J. Fluid Mech.* **258**, 245–254.

Wang, L. & Peters, N. 2006 The length-scale distribution function of the distance between extremal points in passive scalar turbulence. *J. Fluid. Mech.* **554**, 457–475.

Warhaft, Z. 2000 Passive scalars in turbulent flows. *Annu. Rev. Fluid Mech.* **32**, 203–240.

Wu, J-Z., Zhou, Y. & Fan, M. 1999 A note on kinetic energy, dissipation and enstrophy. *Phys. Fluids* **11**, 503–505.

Ying, L., Demanet, L. & Candès, E. J. 2005 3D Discrete curvelet transform. *Tech. Rep*. Applied and Computational Mathematics, California Institute of Technology.

Zaharia, T. & Prêteux, F. 2001 3D shape-based retrieval within the MPEG-7 framework. In *Proc. SPIE Volume 4304: Nonlinear Image Processing and Pattern Analysis XII*, 133–145.

Zeff, B. W., Lanterman, D. D., McAllister, R., Roy, R., Kostelich, E. J. & Lathrop, D. P. 2003 Measuring intense rotation and dissipation in turbulent flows. *Nature* **421**, 146–149.

Zelnik-Manor, L. & Perona, P. 2005 Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems (NIPS'04)*, 1601–1608. MIT Press.

Subject Index